CRWR Online Report 09-02

Hypoxia modeling in Corpus Christi Bay using a hydrologic information system

by

Ernest Sin Chit To, B.S., M.S.

David R. Maidment, Ph.D.

May, 2009.

CENTER FOR RESEARCH IN WATER RESOURCES

Bureau of Engineering Research

The University of Texas at Austin

J.J. Pickle Research Campus Austin, TX 78712-4497

This document is available online via World Wide Web at

http://www.crwr.utexas.edu/online.shtml

The Dissertation Committee for Sin Chit To Certifies that this is the approved version of the following dissertation:

Hypoxia modeling in Corpus Christi Bay using a hydrologic information system

Committee:

David R. Maidment, Supervisor

Robert Gilbert

Ben Hodges

Barbara Minsker

Paul Montagna

Hypoxia modeling in Corpus Christi Bay using a hydrologic information system

by

Sin Chit To, B.S. M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2009

Acknowledgements

I would like to thank my advisor, Dr. Maidment, and the rest of my dissertation committee members, Drs. Bob Gilbert, Ben Hodges, Barbara Minsker and Paul Montagna for their guidance, support and insights for my research.

Hypoxia modeling in Corpus Christi Bay using a hydrologic information system

Publication No._____________

Sin Chit To, Ph.D.

The University of Texas at Austin, 2009

Supervisor: David R. Maidment

Hypoxia is frequently detected during summer in Corpus Christi Bay, Texas, and causes significant harm to benthic organism population and diversity. Hypoxia is associated with the density stratification in the Bay but the cause of stratification is uncertain. To support the study of hypoxia and stratification, a cyberinfrastructure based on the CUAHSI (Consortium of Universities for the Advancement of Hydrologic Science, Inc) Hydrologic Information System (HIS) is implemented. HIS unites the sensor networks in the Bay by providing a standard data language and protocol for transferring data. Thus hypoxia-related data from multiple sources can be compiled into a structured database.

In Corpus Christi Bay, salinity data collected from many locations and times are synthesized into a three-dimensional space-time continuum using geostatistical methods. The three dimensions are the depth, the distance along a transect line, and time. The kriged salinity concentration in space and time illuminates the pattern of movement of a

saline gravity current along the bottom of the Bay. The travel time of a gravity current in the Bay is estimated to be on the order of one week and the speed is on the order of 1 km per day. Statistical study of high-resolution wind data shows that the stratification pattern in the Bay is related to the occurrence of strong, southeasterly winds in the 5 days prior to the observation. This relationship supports the hypothesis that stratification is caused by the wind initiating hypersaline gravity currents which flow from Laguna Madre into Corpus Christi Bay.

An empirical physical hypoxia model is created that tracks the fate and transport of the gravity currents. The model uses wind and water quality data from real-time sensors published by HIS to predict the extent and duration of hypoxic regions in the Bay. Comparison of model results with historical data from 2005 to 2008 shows that wind-driven gravity currents can explain the spatially heterogeneous patterns of hypoxic zones in Corpus Christi Bay.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

1.1 Hypoxia as an environmental problem

Hypoxia in aquatic systems refers to waters where the dissolved oxygen concentration is below 2 mg/L (Dauer et al., 1992). Most organisms avoid, or become physiologically stressed, in waters with oxygen below this concentration (Diaz and Rosenberg, 1995). Hypoxia can also kill marine organisms which cannot escape the low-oxygen water, affecting commercial harvests and the health of impacted ecosystems. While hypoxia can occur naturally, it is also a symptom of environments stressed by human impact such as from excess nutrient enrichment from point and non-point sources. Over half of U.S. estuaries now experience natural or human-induced hypoxic conditions at some time each year and evidence suggests that the frequency and duration of hypoxic events have increased over the last few decades (NOAA, 2007). According to Diaz and Rosenberg (2008), hypoxia problems are increasing world-wide, so providing and understanding of the mechanisms and possible solutions is a vital new problem.

1.2 Hypoxia in Corpus Christi Bay

Hypoxia in Corpus Christi Bay, Texas was first documented in 1988 (Montagna and Kalke, 1992) and later observed every summer (Martin and Montagna, 1995; Applebaum *et al.* 2005). Hypoxia in Corpus Christi Bay results in about a ten-fold reduction in benthic standing stock and diversity.

Unlike other systems along the Gulf of Mexico, e.g. the Louisiana coast, a linkage between eutrophication and hypoxia has not been established in Corpus Christi Bay (Applebaum et al, 2005). Instead, hypoxia is found to be correlated with salinity-induced stratification of the Bay, which occurs in summer when temperature and evaporation are high and precipitation is low (Ritter and Montagna, 1999). Stratification can cause hypoxia by reducing vertical turbulent mixing of heat, momentum, mass and constituents (Ralston and Stacey, 2005; Armenio and Sarker, 2002), and therefore limit the replenishment of dissolved oxygen in the bottom layer. Over time, benthic demand depletes dissolved oxygen to hypoxic levels. As a result, unlike many other coastal and estuarine systems, hypoxia in the Bay is thought to be naturally occurring and driven by stratification, instead of being man-made and driven by nutrient loadings.

Figure 1.1 shows the areas where stratification-induced hypoxia has been observed in Corpus Christi Bay. They were observed by:

1. Dr. Paul Montagna, a marine biologist from the Harte Research Institute of Texas A&M University, Corpus Christi. The region discovered by Montagna is highlighted in brown.
2. Dr. Ben Hodges, an environmental engineer at University of Texas at Austin. The region discovered by Hodges is highlighted in orange.

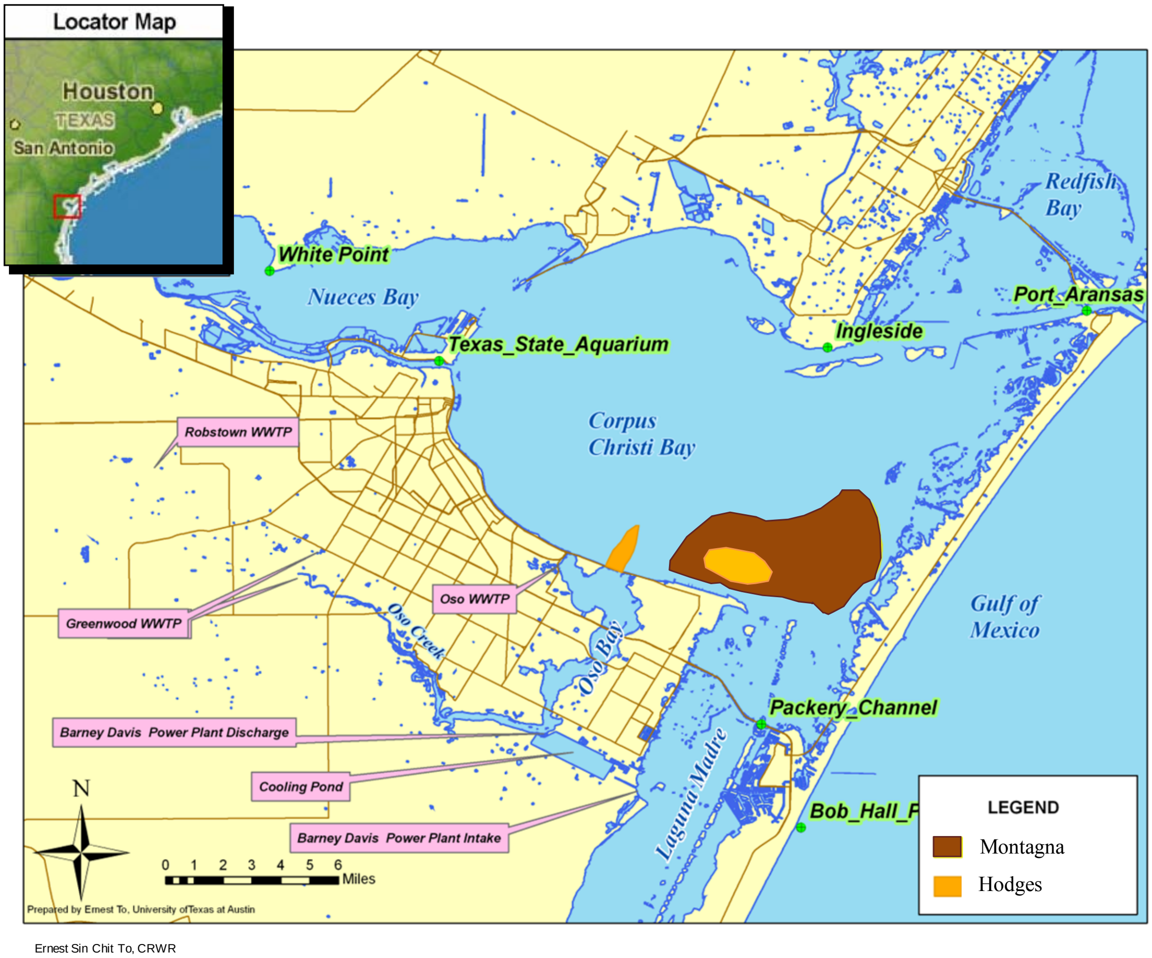


Figure 1.1. Hypoxia in Corpus Christi Bay

1.3 Investigating the hypoxia problem

Because the exact cause and occurrence of stratification and hypoxia in the Bay are not completely known, a team of biologists, engineers and computer scientists from four universities has been assembled to study the problem. To enhance collaboration among its members, this team implemented a cyberinfrastructure in its investigation. A cyberinfrastructure is a result of recent advances in information technology. It is defined as an integrated system of computers (ranging from laptops to supercomputers), data storage systems, advanced instruments, visualization environments, and people that are all linked by high speed networks to foster innovation and discoveries not otherwise tenable. The cyberinfrastructure used in Corpus Christi Bay is the CUAHSI (the Consortium of Universities for the Advancement of Hydrologic Science, Inc) HIS (Hydrologic Information System). HIS was released in August 2007 and is the first national-level cyberinfrastructure for the hydrologic sciences. The goal of HIS is to unify and streamline the flow of environmental data among governmental and research organizations. HIS is a system of web-interactive programs, protocols and computer servers that allow participators to

1) publish data on the internet,

2) discover relevant data from other sources and

3) ingest data into modeling or analytical workflows.

By adopting HIS, researchers provide their computer systems with the ability to interact, on a machine-to-machine basis, with the systems of their peers, thereby allowing them to become part of the overall cyberinfrastructure. HIS benefits the overall scientific community by enlarging the common pool of data and resources through its cyberinfrastructure. In Corpus Christi Bay, sensor networks from various state agencies and academic institutions have measured environmental data in since the 1970s. The implementation of HIS enables data from the various environmental sensor networks in Corpus Christi Bay to be published and accessed in a unified manner.

1.4 Modeling the hypoxia problem

Although the potential of cyberinfrastructure is much touted, the process by which information technology leads to scientific discovery is not well understood. Since its release, HIS has been implemented on more than ten study areas (also known as "testbeds") across the nation to evaluate its effectiveness in the advancement of hydrologic science. These testbeds cover a variety of systems, ranging from rivers, estuaries to oceans. Each testbed has a unique environmental phenomenon that is being researched by a team of scientists. Corpus Christi Bay is one of these testbeds.

Apart from investigating the hypoxia problem in Corpus Christi Bay, the goal of this research is to use the Corpus Christi Bay testbed to conceive and implement a framework for utilizing hydrologic information systems, and thus cyberinfrastructures, for advancing science. This framework shows how HIS can be harnessed to:

1. generate insights into an environmental process (e.g. hypoxia) that is previously not well-understood;
2. support the development of a model of the environmental process; and,
3. support the execution of the model to generate predictions over space and time.

By demonstrating that cyberinfrastructure can directly contribute to scientific discovery, this research proves they are a critical component in future collaborative research. The framework is described in the next section.

1.5 Framework for harnessing HIS for scientific discovery

The framework developed for investigating hypoxia in Corpus Christi Bay addresses the question, "Now that I have HIS, what do I do with it?" It is described in the steps below and illustrated in Figures 1.2 and 1.3.

1. Data compilation

   First, the relevant data are harvested from various sources into a local database. Because different sources follow different naming conventions for labeling their data, semantic mediation is performed to sort out similar and dissimilar variables.

2. Data analysis

   After compilation, the local database exhibits two main characteristics:

   1). It contains a collection of data that are scattered over space and time.

   2). It contains a collection of environmental variables that describe different environmental factors.

   For 1), the scattered data can be synthesized spatially and temporally into a continuous space-time volume. Cross-sections of can be sliced from the space-time volume to visualize how a particular variable changes over space and time. Spatial and temporal trends, such as velocity, can be observed from such visualizations.

   For 2), the relationships among environmental factors can be observed by performing hypothesis-testing of associated variables. For instance, cause-and-effect relationships can be elicited by appropriately applying correlation or partial correlation tests. Needless to say, hypothesis posed need to be grounded on sound scientific principles in order to extract meaningful information.

3. Knowledge integration

Knowledge of trends and relationships that are gained from data analysis can be incorporated into a parsimonious model of the environmental process. A parsimonious model is one that is very careful or economical in its use of model parameters, and thus requiring only the available input data. The parsimonious model, once calibrated and validated, is a reflection of how much we understand about an environmental system.

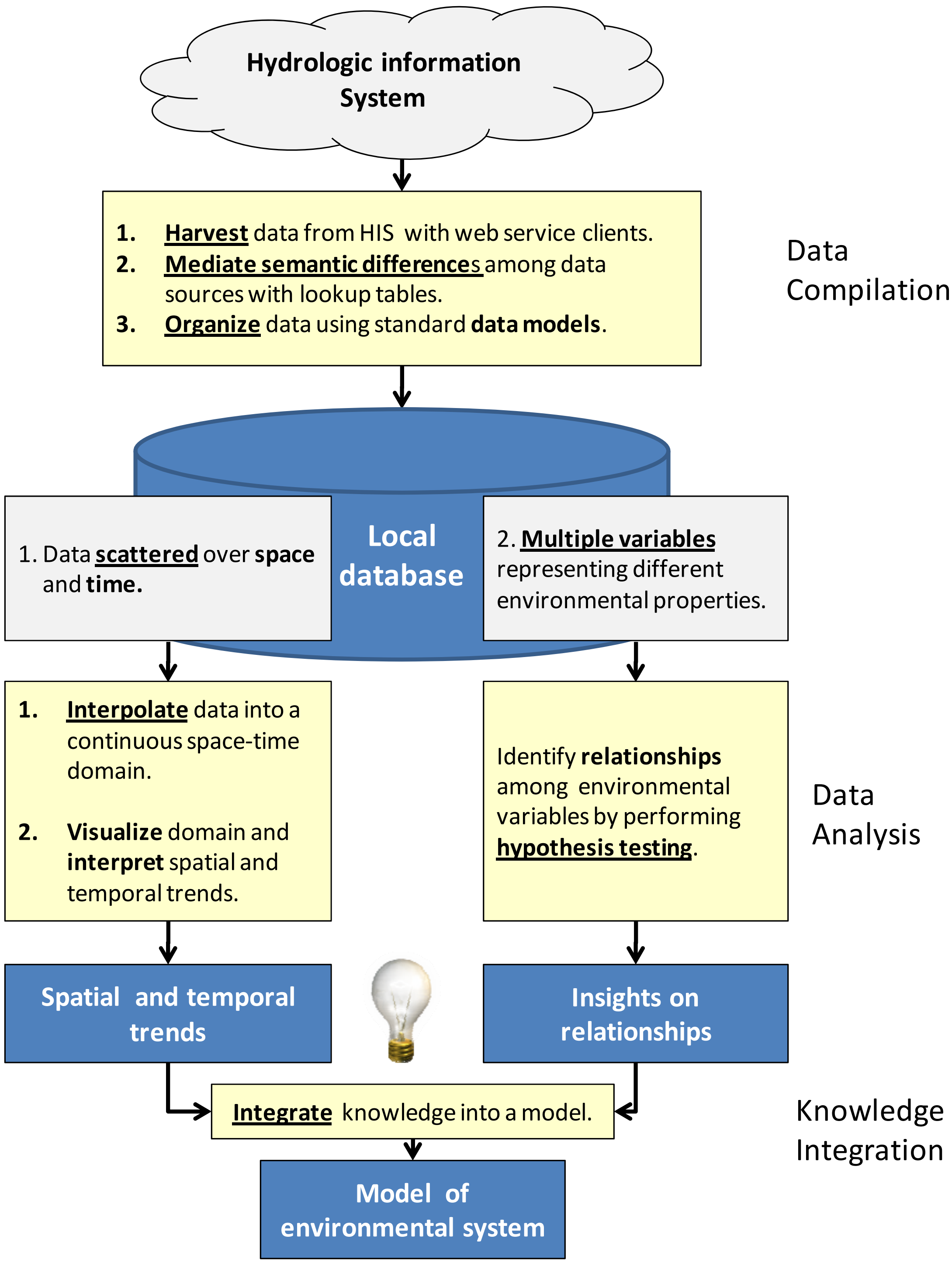


Figure 1.2. Framework for harnessing hydrologic information system for scientific discovery.

4. Simulation

The model can be connected to data sources via HIS generate predictions and estimations of the environment system.

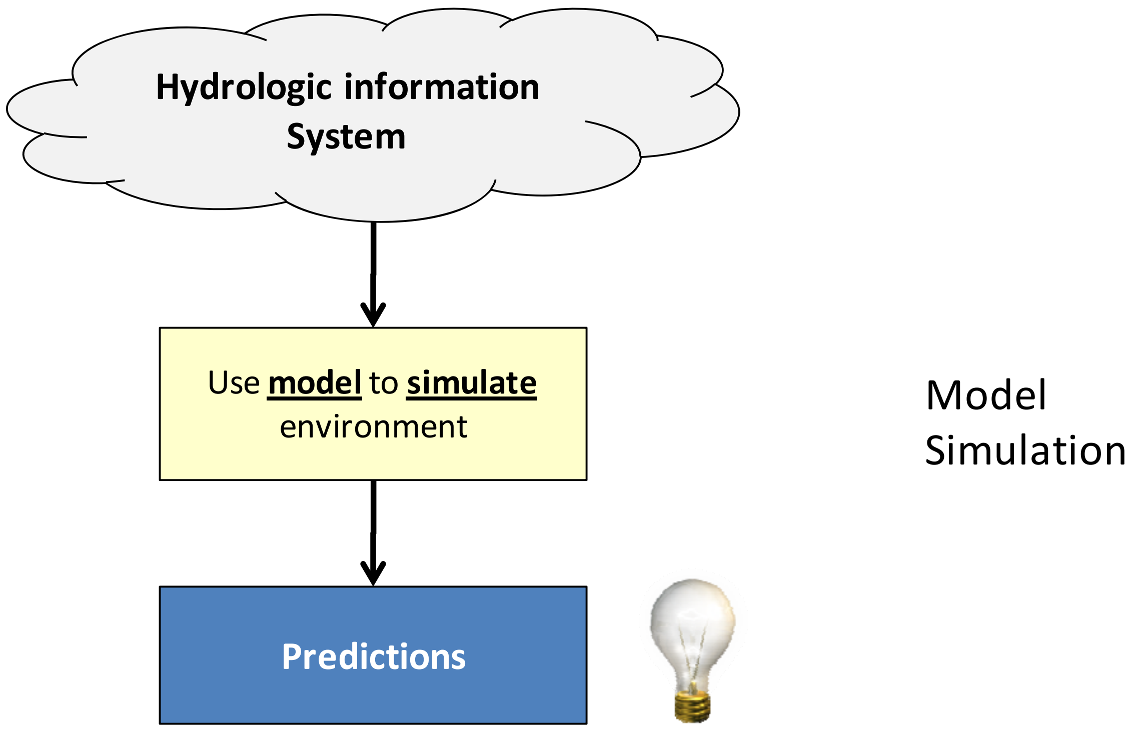


Figure 1.3. Using HIS to support execution of environmental models.

1.6 Research Questions

Several technologies need to be developed in order for the above process to be possible. The first hurdle is to figure out how data compilation can be performed. More specifically, the technology or protocol for assembling data from multiple sensor network needs to be developed. So the first research question is:

1. ***How can data be assembled from a service-oriented architecture of environmental sensor networks to describe the properties of a water domain in space and time?***

The next step is to understand how data from multiple sensor networks can be synthesized. In other words, how can time series data from various locations be integrated into a continuous space-time domain to describe a variable across space and time. Hence the second research question is:

2. ***How can data collected at different locations, times, spatial resolutions and temporal frequencies be synthesized to provide a continual description of an environmental variable over space and time?***

The third step is to extract scientific knowledge from the data. Prior knowledge of the environmental system is used to frame hypotheses about the environmental phenomenon. By developing experiments around HIS data, the hypotheses are tested and insights about the underlying mechanisms are gained. Therefore the third question is:

3. ***How can data about different environmental variables be used to generate insight about underlying mechanisms about a given environmental phenomenon?***

The last step is to integrate space-time trends and knowledge gained from hypotheses-testing into useful models. These models take advantage of HIS data streams to generate useful predictions. Therefore the final question is:

4. ***How can new models be designed around hydrologic information systems to make predictions about a given environmental phenomenon? How do models explain the hypoxia patterns in southeast Corpus Christi Bay?***

1.7 Organization of dissertation

This dissertation consists of a series of four related papers to address each of the three research questions. The papers are included in chapters 2 to 5. The titles of the chapters are listed in Table 1.1 below.

Table 1.1. Titles of paper chapters in dissertation.

Chapter	**Chapter title**
2	*Harvesting data from hydrological information systems for environmental research in Corpus Christi Bay*
3	*Using space-time interpolation to characterize movement of gravity currents in Corpus Christi Bay*
4	*Effects of wind on salinity stratification in southeast Corpus Christi Bay.*
5	*Modeling the effects of wind on hypoxia in southeast Corpus Christi Bay.*

Chapter 6 discusses how the four research questions are answered and summarizes the contributions of this research to science. It also describes how improvements can be made in the future.

CHAPTER 2: HARVESTING DATA FROM HYDROLOGIC INFORMATION SYSTEMS FOR ENVIRONMENTAL RESEARCH IN CORPUS CHRISTI BAY

By Sin Chit To, Timothy L. Whiteaker and David R. Maidment

2.1 Abstract

Hypoxia is frequently detected during summer in Corpus Christi Bay and causes significant harm to benthic organism population and diversity. Several governmental agencies and research institutes have monitored the Bay since the 1970s, thus providing a wealth of data available for hypoxia research. However, access to the data is often impeded by the disparate formats and methods by which the Bay data are stored and published. To unify the data sources, a cyberinfrastructure based on the CUAHSI (Consortium of Universities for the Advancement of Hydrologic Science, Inc) Hydrologic Information System (HIS) is deployed. HIS employs a web service oriented architecture to provide a common interface for users to query and download data from the different sensor networks. Two kinds of CUAHSI data web services are implemented in Corpus Christi Bay. Generic OD web services are implemented to handle data sources that are updated infrequently while the more complex Hybrid web services are implemented on data sources that are updated in real time. The consolidation of data is achieved by using a software tool called HydroGET. HydroGET works in a geographic

information system environment and can 1) batch process data requests for multiple sites, variables and time periods 2) store and organize data using the *Arc Hydro* data model; and, 3) mediate the differences in terminologies among heterogeneous data sources using a MySelect table.

HydroGET is applied successfully to the sensor networks in Corpus Christi Bay to harvest hypoxia-related data. The organization of the data within the Arc Hydro data model supports the clear presentation of spatial and temporal patterns of the harvested data using the Datacube diagram.

2.2 Introduction

BACKGROUND

Hypoxia is detected episodically during the summer in Corpus Christi Bay in south Texas and is known to cause a dramatic reduction in benthic organisms in the Bay (Coopersmith, et. al., 2007). Corpus Christi Bay is an urban estuary that is located near the Gulf of Mexico (see Figure 2.1) next to the city of Corpus Christi.

Hypoxia in the Bay was first documented in 1988 (Montagna and Kalke, 1992) and later observed every summer (Martin and Montagna, 1995; Applebaum *et al.* 2005). Hypoxia is associated with salinity-induced stratification of the Bay, which occurs in summer when temperature and evaporation are high and precipitation is low (Ritter and Montagna, 1999). Studies by Russell and Montagna (2007) have shown that stratification is correlated with environmental variables like wind, salinity, dissolved oxygen and temperature and depth.

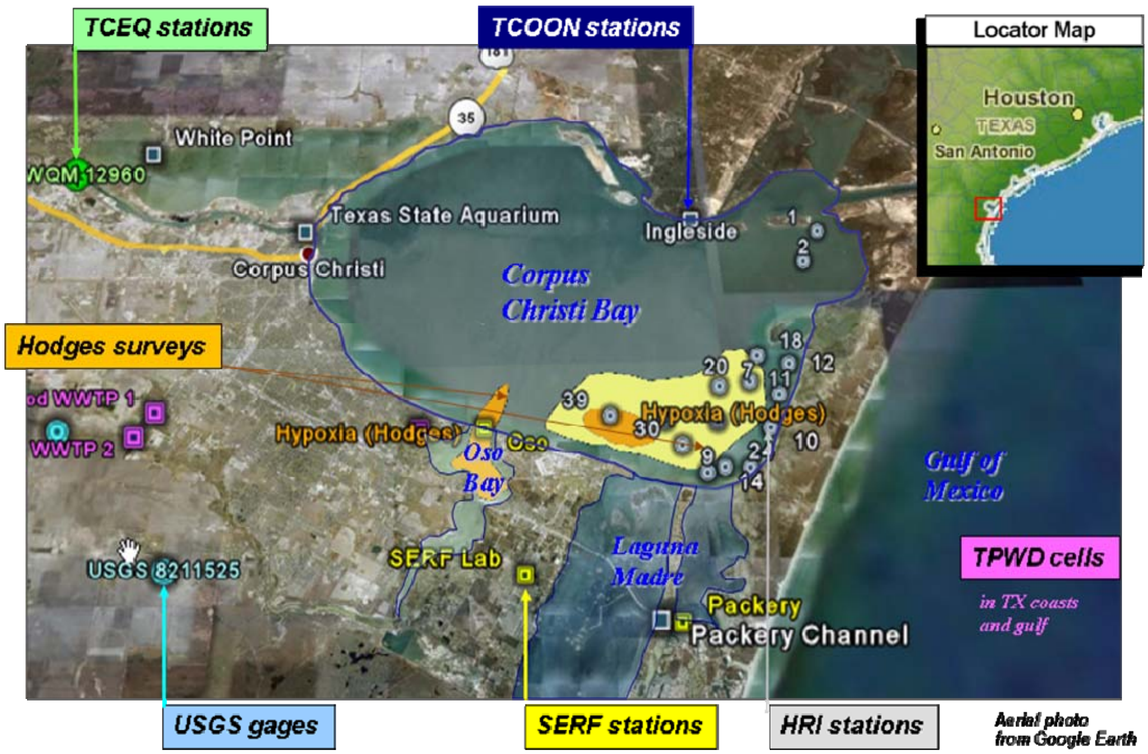


Figure 2.1. Sensor networks in Corpus Christi Bay (To, 2008).

Because the exact cause of stratification and hypoxia in the Bay is not known, a study is being conducted by a team of scientists and engineers from four universities (the University of Illinois at Urbana-Champaign, Texas A&M University College Station, University of Texas at Austin and Texas A&M University Corpus Christi) to understand the underlying mechanisms. The foremost step in the investigation is to compile available hypoxia-related data to support the study. This paper describes how modern information technology is used to perform the data compilation.

DATA SOURCES IN CORPUS CHRISTI BAY

Since the 1970s, several governmental agencies and research institutes have maintained sensor networks in the Bay to measure environmental variables such as salinity, dissolved oxygen and wind speed. These networks include the HRI (Harte Research Institute) network, TWDB (operated by Dr. Ben Hodges of the University of Texas at Austin and funded by Texas Water Development Board) surveys, SERF (Shoreline Environmental Research Facility) network, TCEQ (Texas Commission on Environmental Quality) network, TCOON (Texas Coastal Ocean Observation Network) network, TPWD (Texas Parks and Wildlife Department) network and the USGS (United States Geological Survey) network. Figure 2.1 shows the locations of the networks. Table 2.1 provides a brief description of each network and the data they collect.

Table 2.1. Sensor networks in Corpus Christi Bay

	Name of network	**Description**	**Data**
1.	HRI	Monitoring stations maintained by Dr. Paul Montagna of the Harte Research Institute	Collects grab samples of dissolved oxygen, salinity, temperature and other water quality data on a bi-weekly to monthly basis during summer months.
2.	TWDB surveys (by Dr. Hodges)	Plume tracking studies performed by Dr. Ben Hodges of the University of Texas at Austin	Collects grab water quality samples at high temporal (12 hour) and spatial resolutions to track the movement of hypersaline waters from Oso Bay and Laguna Madre (Hodges, et al., 2008, Brower, *et al.* 2007).
3	SERF	Shoreline Environmental Research Facility	Collects continuous water quality data and surface current velocities.
4	TCEQ	Texas Commission on Environmental Quality	Collects water quality data in water bodies in Texas
5	TCOON	Texas Coastal Ocean Observation Network	Collects continuous wind and tide data along the coast of Texas.
6	TPWD	Texas Parks and Wildlife Department	Collects grab water quality samples as part of its biological sampling program along the coast and estuaries of Texas. Sampling is performed using a 1 minute x 1 minute grid over the Texas coast.
7	USGS	United States Geological Survey	Collects hydrodynamic and water quality data for streams and rivers.

CHALLENGES IN INTEGRATING SENSOR DATA

Despite the presence of several sensor networks, access to the data is impeded by the disparate formats and methods used by the data sources to publish their data. The methods used by the data sources (prior to this research) are summarized as follows:

1. SERF, TCOON and USGS publish data on their web sites almost immediately after collection. Users can access the data using online forms. However the formats of the forms are different for each data sources. In addition, the formats of the data returned from the websites are not the same.
2. TCEQ publishes data on a periodic basis on its website. Unfortunately the web site cannot handle queries for specific sites or variables and only facilitates bulk downloads of data for individual water quality segments.
3. TPWD, TWDB, HRI store data within in-house databases and publish them in scientific reports. Users had to manually request the data from the source in order to get them.

The differences in publication methods make data compilation a tedious task. Not only is significant time consumed in requesting the data, but also in reformatting the data and consolidating them into a database. To tackle the challenges posed by disparate data publication methods, a Hydrologic Information System (HIS) is implemented to unify access to the data sources.

2.3 What is a hydrologic information system (HIS)?

The Hydrologic Information System (HIS) is a cyber-infrastructure that is designed by CUAHSI's (Consortium of Universities for the Advancement of Hydrologic Science, Inc) Hydrologic Information System (HIS) Project (CUAHSI, 2007). The immediate goal of HIS is to unify access to the nation's water information by creating a set of tools and communication protocols for publishing and accessing environmental data. The advantages of the system are three-fold:

1. data users can search for environmental data across a multitude of sources instead of one source at a time;
2. computers of data users can directly ingest environmental data from data servers without the need for human intervention;
3. data providers can publish data in a widely-accepted format – thus obviating the need to cater to individual data requests.

HIS is built upon a distributed network of computer servers that are linked together by the internet. These computer servers, known as HIS servers, host a family of web services, tools and applications. Researchers can easily setup HIS servers following the HIS manual published by CUAHSI (CUAHSI, 2007). These various HIS technologies function together as an integrated system, allowing hydrologic data to be stored, found, accessed, interpreted and analyzed.

WEB SERVICES: THE BUILDING BLOCKS OF HIS

CUAHSI-HIS is built using a service-oriented architecture (SOA) in which the fundamental building block is the web service. In simple terms, a web service is a computer application that is executed remotely over the internet. Web services provide a variety of functions, such as computation, data retrieval and manipulation. Web services allow users to access technologies without knowledge of, expertise with, or control over the infrastructure that supports them. Users can create scientific workflows that rely on existing web services to perform analyses.

CUAHSI's data web services act as intermediaries between data sources and users. Typically, one web service is built for each data source. The web service provides a standard interface for querying data from the data source. To get data, the user submits a request to the web service that contains standard parameters that include 1) the location of interest, 2) the variable desired and 3) the start and end data of the period of interest. The web service translates these parameters to a form understandable by the computer system of the target data source. When the data source returns the data, the web service transforms the data into a common language known as WaterML (Goodall, et al., 2008) and then sends them to the user. This process is abbreviated as ETL (Extract – Transform – Load) (see Figure 2.2). The concept, design and function of CUAHSI data web services are explained in detail in the paper by Goodall, et al., 2008.

Apart from queries for getting the data, CUAHSI data web services can also support queries to the catalog of the data. This means the user can discover what sites and variables are measured by a data source. The user can use this information to determine what data to download. The functions that support queries to the data are known as data delivery methods. The functions that support queries to the data catalog are known as data discovery methods.

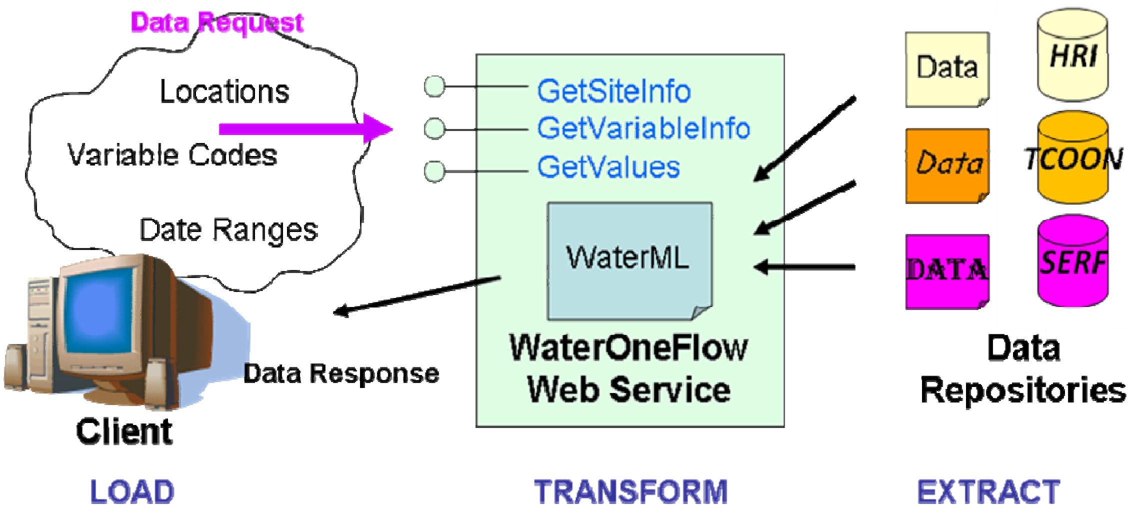


Figure 2.2. How CUAHSI data web services work.

STATUS OF IMPLEMENTATION OF HIS IN CORPUS CHRISTI BAY

CUAHSI data web services have been implemented for six of the seven sensor networks in Corpus Christi Bay. The web addresses of the web services are shown in Table 2.1.

Table 2.1. Web service locations of Corpus Christi Bay environmental sensor networks

	Name of network	**Web service address (WSDL)**
1	HRI	http://ccbay.tamucc.edu/CCBayODWS/cuahsi_1_0.asmx?WSDL
2	SERF	http://his.crwr.utexas.edu/serf/serf.asmx?wsdl
3	TCEQ	http://his.crwr.utexas.edu/TRACS/cuahsi_1_0.asmx?WSDL
4	TCOON	http://his.crwr.utexas.edu/tcoonts/tcoon.asmx?wsdl
5	TPWD	http://his.crwr.utexas.edu/TPWDCoast/cuahsi_1_0.asmx?wsdl
6	USGS	http://water.sdsc.edu/WaterOneFlow/NWIS/DailyValues.asmx?wsdl

The method of their implementation is summarized as follows:

1. For TPWD, TCEQ, and HRI, the in-house databases mentioned previously are transferred to computer servers where they are migrated wholesale to a standard database schema known as the ODM (Observations Data Model). Generic web services designed by CUAHSI are installed over these new databases to make their data accessible on the web. The wholesale migration method is a simple implementation method that requires little programming skills. It also automatically generates catalog of the data during the implementation. This data catalog supports data discovery methods of the CUAHSI data web. However, the migration process can be tedious because of the need to map the original database schema to the ODM schema. Therefore it is more suitable for databases

that are updated less frequently (i.e. less than once a month). For this reason, this method is used for the TPWD, TCEQ and HRI sensor networks.

2. For SERF, TCOON and USGS, a different kind of web service called a hybrid web service is used. The wholesale migration method is not applicable for these networks because their databases are continuously updated by real-time data. Fortunately, the data for these sources are published on the web, so data delivery requests are satisfied by mediating data transfers between the user and the source web sites. The mediation is performed by functions called "webscrapers" that extract and translate contents from web pages into WaterML. Data discovery methods for TCOON and USGS are supported by a data catalog created by a partial migration of the metadata into ODM and followed by installation of generic OD web services. The process is less simple and is out of the scope of this paper. On the whole, because of the need to work with real-time data, hybrid services are more complicated to deploy than generic OD web services.

HOW DO WE CONSOLIDATE DATA FROM HIS?

The implementation of HIS in Corpus Christi Bay has made available a large collection of hypoxia-related data. The next challenge is to develop a methodology to consolidate data from multiple sources into a structured database to support scientific research. Different data sources have varying naming conventions for naming sites and variables that need to be mediated before they can be gathered into one database. Therefore, the methodology needs to address the following three questions:

How can data from multiple web services be harvested?

This question calls for the ability batch process multiple web service requests. HydroGET, a web service client for ArcGIS, is tool developed to harvest data for multiple sites and variables across different networks.

How should the harvested data be stored?

This question calls for a data model that is robust enough to store heterogeneous data from multiple sources. The Arc Hydro data model (Maidment, 2002) contains a highly versatile time series component that is designed to handle point observations in space-and-time.

How should harvesting be managed and semantic differences be mediated?

This question calls for a data tagging system that groups "like" variables together and differentiates "unlike" variables so that data can be properly organized in a database. The MySelect table is a lookup table that tags variables from different sources with internal identifiers so that differences and similarities can be sorted out.

2.4 Methodology

A methodology is developed to address the three research questions. It involves three technological components: CUAHSI's HydroGET (To, 2008), the Arc Hydro data model (Maidment, 2002) and a table schema called MySelect. It is described in the following.

USING CUAHSI'S HYDROGET TO PROCESSES MULTIPLE WEB SERVICE REQUESTS

CUAHSI's HydroGET tool is designed to assemble data from multiple web services into one database. HydroGET (Hydrologic GIS Extraction Tool) is a Geographic Information System (GIS) application and runs within ESRI™'s ArcMap environment. It interacts with site locations on the map (represented as feature classes) and downloads data for them. Figure 2.3 shows the interface for HydroGET. The locations where data are desired are specified by selecting a feature class from the ArcMap's table of contents. The desired variable (e.g. precipitation, streamflow) is chosen from one of HydroGET's five tabs: Atmosphere, Surface, Subsurface, Custom (Single Point) and Custom (Multiple Points). Finally, the period of interest and the output location for the data are specified in the interface. HydroGET stores the downloaded data into an Arc Hydro geodatabase (Maidment, 2002). A description of the Arc Hydro geodatabase is discussed in the next section.

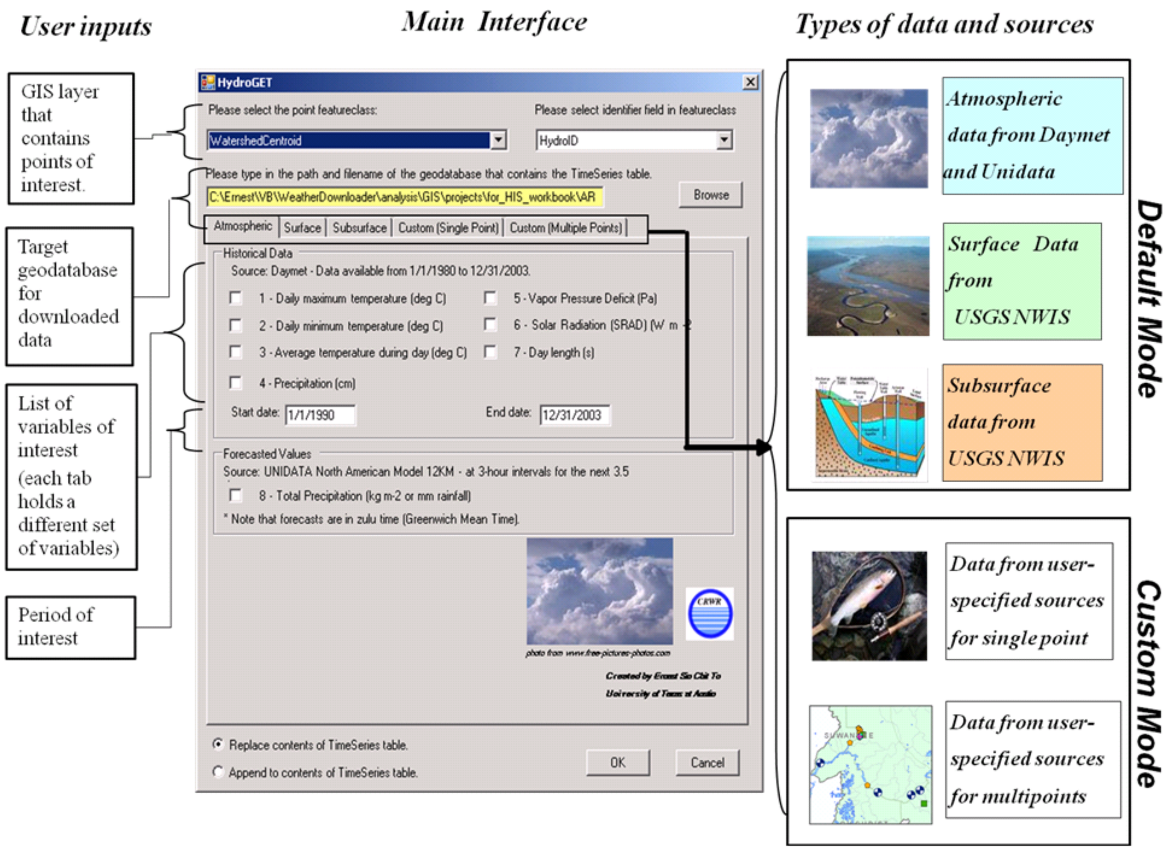


Figure 2.3. The HydroGET interface (To, *et al.* 2007)

Once executed, HydroGET cycles through each location within the selected feature class, and downloads the desired data through CUAHSI data web services. Time series data are stored into the TimeSeries table of an ArcHydro geodatabase (see Figure 2.4), while information regarding the data variable are stored into the TSType table.

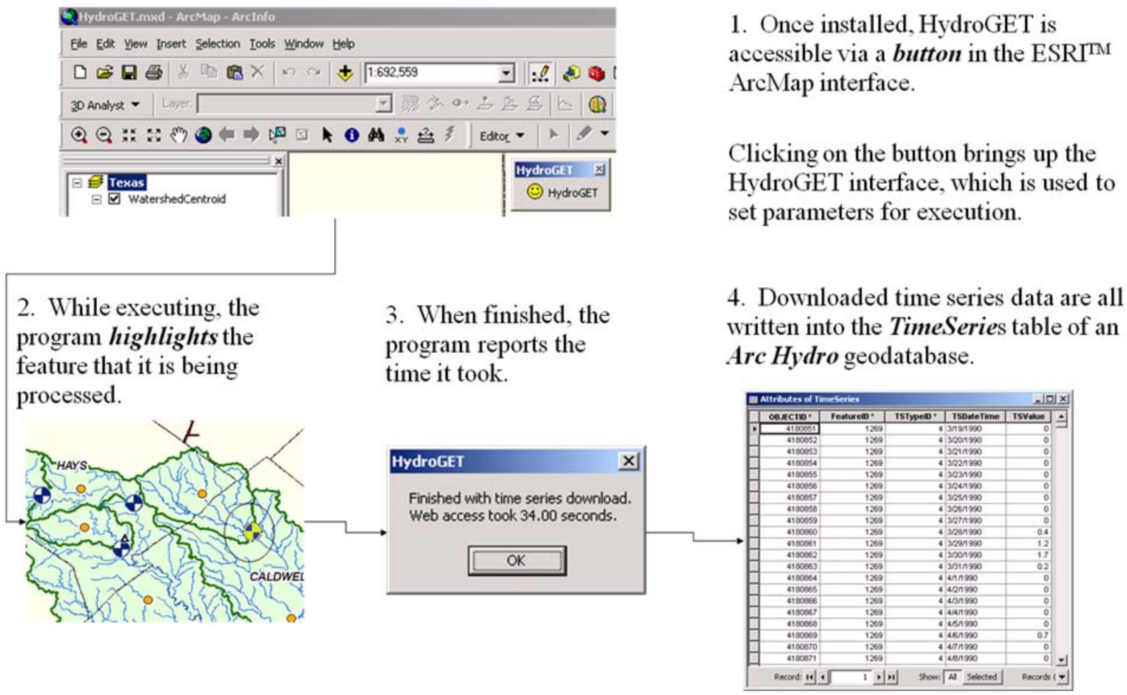


Figure 2.4. The operation of HydroGET.

A detailed tutorial of the HydroGET application, associated installation files and source code are available from the CUAHSI-HIS web site at: http://his.cuahsi.org/hydroget.html (To, *et al.* 2008).

USING THE ARC HYDRO DATA MODEL TO STORE HETEROGENEOUS TIME SERIES DATA

HydroGET stores the downloaded data into an Arc Hydro geodatabase, which is a relational database that implements the Arc Hydro data model (Maidment, 2002). The Arc Hydro data model uses the data cube concept to integrate time series data from heterogeneous data providers. The data cube, also known as a space-time-variable cube

(Goodall, et al., 2008), states that every observation is unique in 1) space, 2) time and 3) variable type (e.g. salinity, oxygen, temperature). In other words, a minimum of these three properties is needed to address a given observation (see Figure 2.5). The Arc Hydro data model utilizes one table to store information for each of the three properties. They are

1) the TimeSeries table, which describes the time and the observed value;
2) the MonitoringPoint table, which describes the location; and,
3) the TSType table, which describes the variable type.

Database relationships link the three tables together to describe a set of observations.

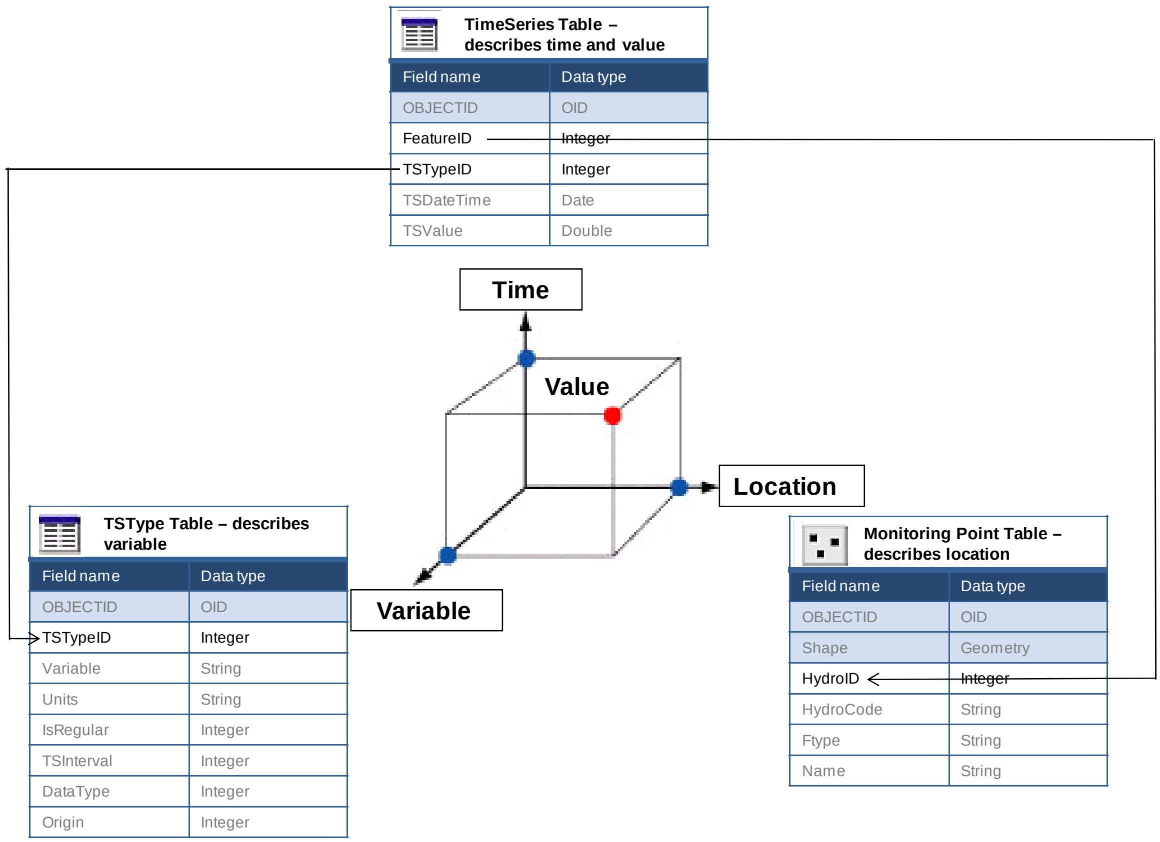


Figure 2.5. Data cube model (Maidment, 2002)

The benefit of using three tables is that it stores the data compactly. Redundancies caused by repeating sites and variables are minimized. The data cube model is often augmented by additional tables to describe ancillary information (e.g. sampling method). For example, the Arc Hydro data model employs additional tables to describe the connectivity of monitoring points with hydrologic features such as streams and lakes.

USING THE MYSELECT TABLE TO MANAGE DATA HARVESTING AND STORAGE

The MySelect table is a special kind of feature class that allows HydroGET to batch process multiple data requests from multiple web services for multiple locations, variables and periods of interest. The attribute table of MySelect contains the essential information needed to query the desired data from CUAHSI data web services. In addition, the table also contains the identifiers for each piece of downloaded data so that they can be properly organized in database.

Structure of MySelect

The structure of MySelect is shown in Figure 2.6. Each row of the table is a data request to a specific web service for one location, variable and time period. The columns of the table can be grouped into three portions:

1. A web service portion contains ***query parameters*** for calling web service.
2. The Arc Hydro portion contains ***identifiers*** to relate web service terminologies for variables and sites to existing database terminologies.
3. The metadata portion is optional. It contains additional information that explains the query parameters. For instance, the site code of 'H1' refers to the site name of 'Hypoxia_1' in the HRI network and the variable code of "DOConcGrab" refers to the variable name of "Dissolved oxygen concentration grab sample". These fields are not read by HydroGET. Their goal is to improve the understandability of the MySelect table.

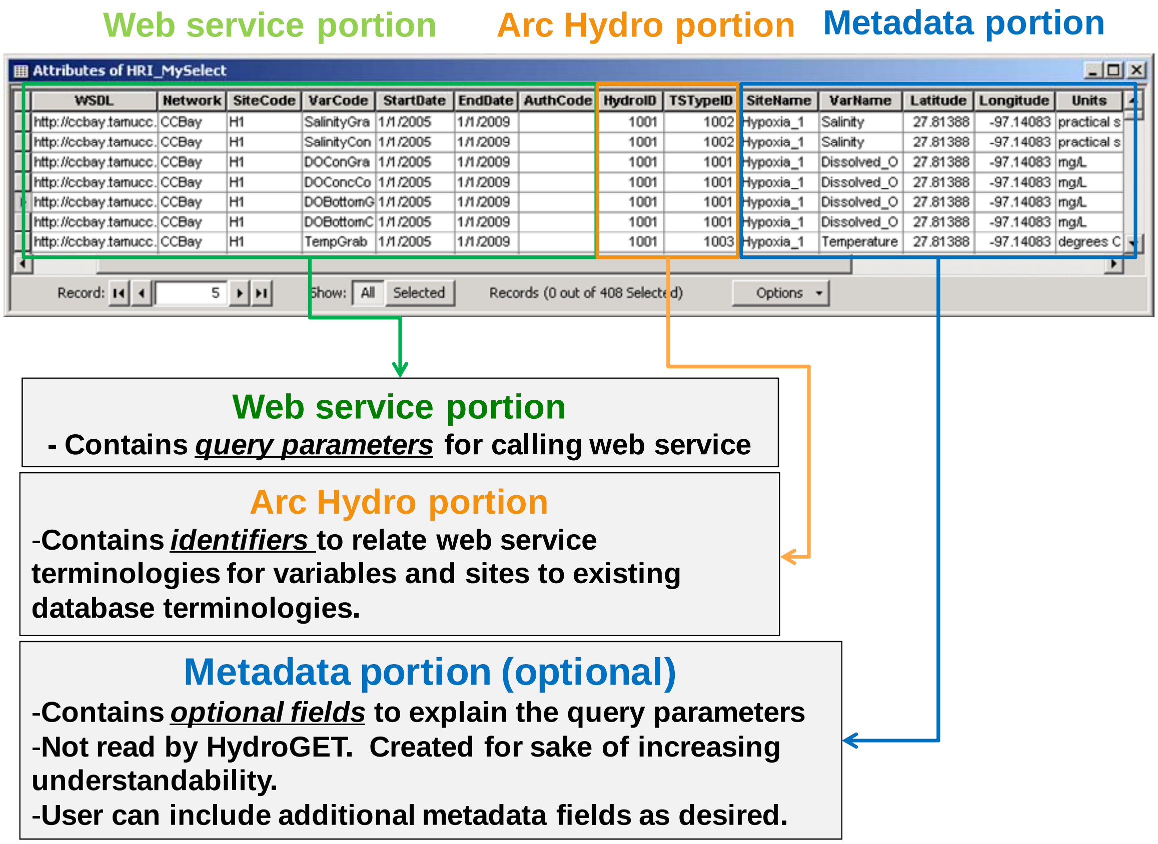


Figure 2.6. Structure of MySelect

The fields in the web service portion of MySelect are described in Table 2.2.

Table 2.2. Explanation of fields in the web service portion of MySelect.

Column	Explanation	Example
WSDL	Web Service Description Language (WSDL). This is the web address of the web service.	http://ccbay.tamucc.edu/CCBayODWS/cuahsi_1_0.asmx?WSDL
NETWORK	Name of the sensor network	CCBay (this is the network name of the HRI network)
SITECODE	Site code of the monitoring point	H1
VARCODE	Variable code for the variable of interest	SalinityGrab (this is the variable code for salinity grab samples)
STARTDATE	Start date for the period of interest	1/1/2005
ENDDATE	End date for the period of interest	1/1/2008
AUTHCODE	Authentication code	

Together, the above set of query parameters can be read as follows:

Get salinity data (variable code: SalinityGrab) data between 1/1/2005 and 1/1/2008 at site H1 from the HRI environmental sensor network (network name: CCBay). The address for the HRI web service is http://ccbay.tamucc.edu/CCBayODws/cuahsi_1_0.asmx?wsdl

The fields in the Arc Hydro portion of MySelect are described in Table 2.3.

Table 2.3. Explanation of fields in the Arc Hydro portion.of MySelect.

HYDROID	Internal ID in the Arc Hydro database for identifying the site where the data was collected.	1001 (a long integer)
TSTYPEID	Internal ID in the Arc Hydro database for identifying the variable type of the data.	1002 (a long integer)

Together, the example information can be interpreted as follows::

Once the salinity data is downloaded, tag the location of the data with the internal HydroID of 1001 and the variable type of the data with the internal TSTypeID of 1002.

The HydroID is a location identifier which is explained in the MonitoringPoint table of the Arc Hydro model while the TSTypeID is a variable identifier explained in the TSType table. New variables do not need to be pre-defined within the Arc Hydro database before downloading the data. When a new TSTypeID is introduced in the MySelect table, HydroGET extracts the related information about the variable from the web service and adds them to the TSType table.

Semantic mediation

The tagging of the data is particularly critical to the consolidation of data. Each sensor networks in Corpus Christi Bay has a different naming convention for environmental variables. Some networks even have two or more names for the same variable. For instance, a total of seven names are given for dissolved oxygen. The HRI network uses 1) “DOConcCon” to name dissolved oxygen that is collected continuously along the depth of the water column, 2) “DOConGrab” to name dissolved oxygen that is collected as grab samples, 3) “DOBottomGrab” to name dissolved oxygen that is collected at the bottom of the Bay as grab samples and 4) “DOBottomCon” to name dissolved oxygen that is collected at the bottom of the Bay as continuous samples. The Hodges, SERF and TPWD networks use 5) “DO”, 6) “oxygen”, and 7) OXY001”, respectively.

To the researcher, all six variables are qualitatively the same. By tagging all these variables with the TSTypeID of 1001 in the MySelect table, they are consolidated as one variable within the Arc Hydro database. Thus, MySelect acts as lookup table for mediating differences in semantics among the different data sources. An illustration is provided in Figure 2.7.

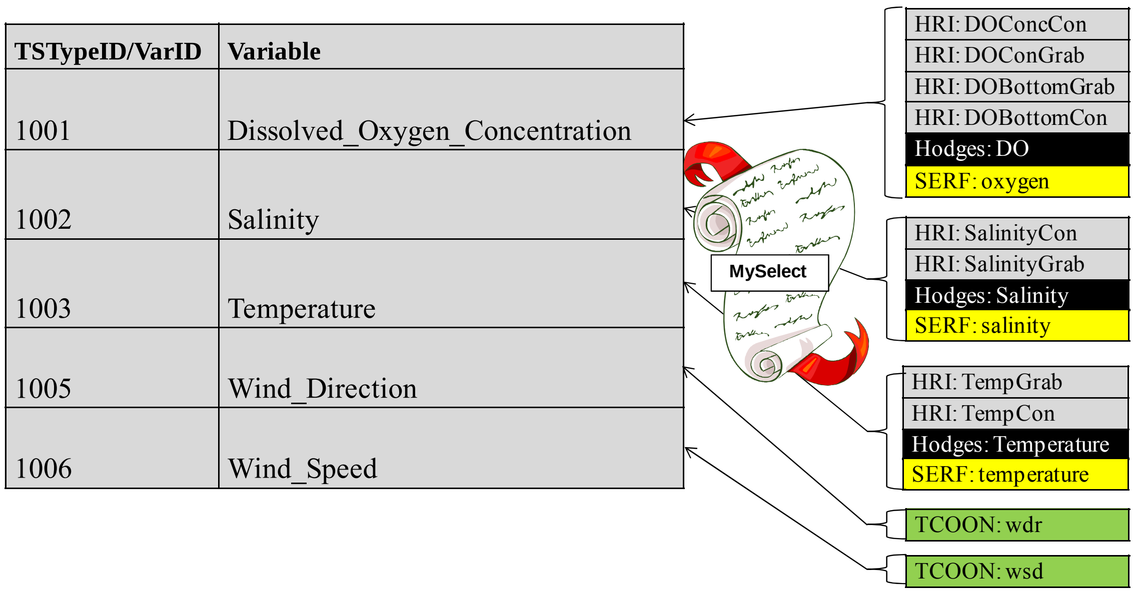
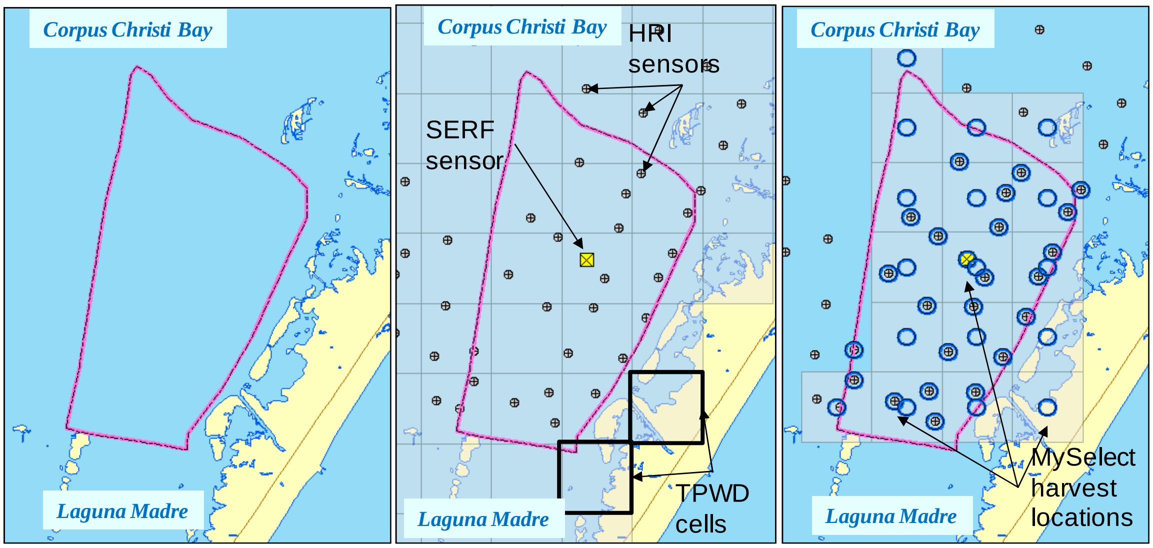


Figure 2.7. Semantic mediation by MySelect

2.5 Application to Corpus Christi Bay

An example of how dissolved oxygen data was harvested in Corpus Christi Bay is presented to illustrate the application of the consolidation methodology. First an area near Laguna Madre at the southeast corner of the Bay is chosen as the study area. Next, the sensors from networks that measure dissolved oxygen are overlaid with the study area. A feature class is created by merging all the sites that intersect the study area into one feature (see Figure 2.8).



1. A study area is selected in Corpus Christi Bay

2. Sensor networks that measure dissolved oxygen are overlaid with the study area.

3. A MySelect feature class is created to harvest dissolved oxygen data from all the sites that intersect the study area.

Figure 2.8 Creating a MySelect featureclass.

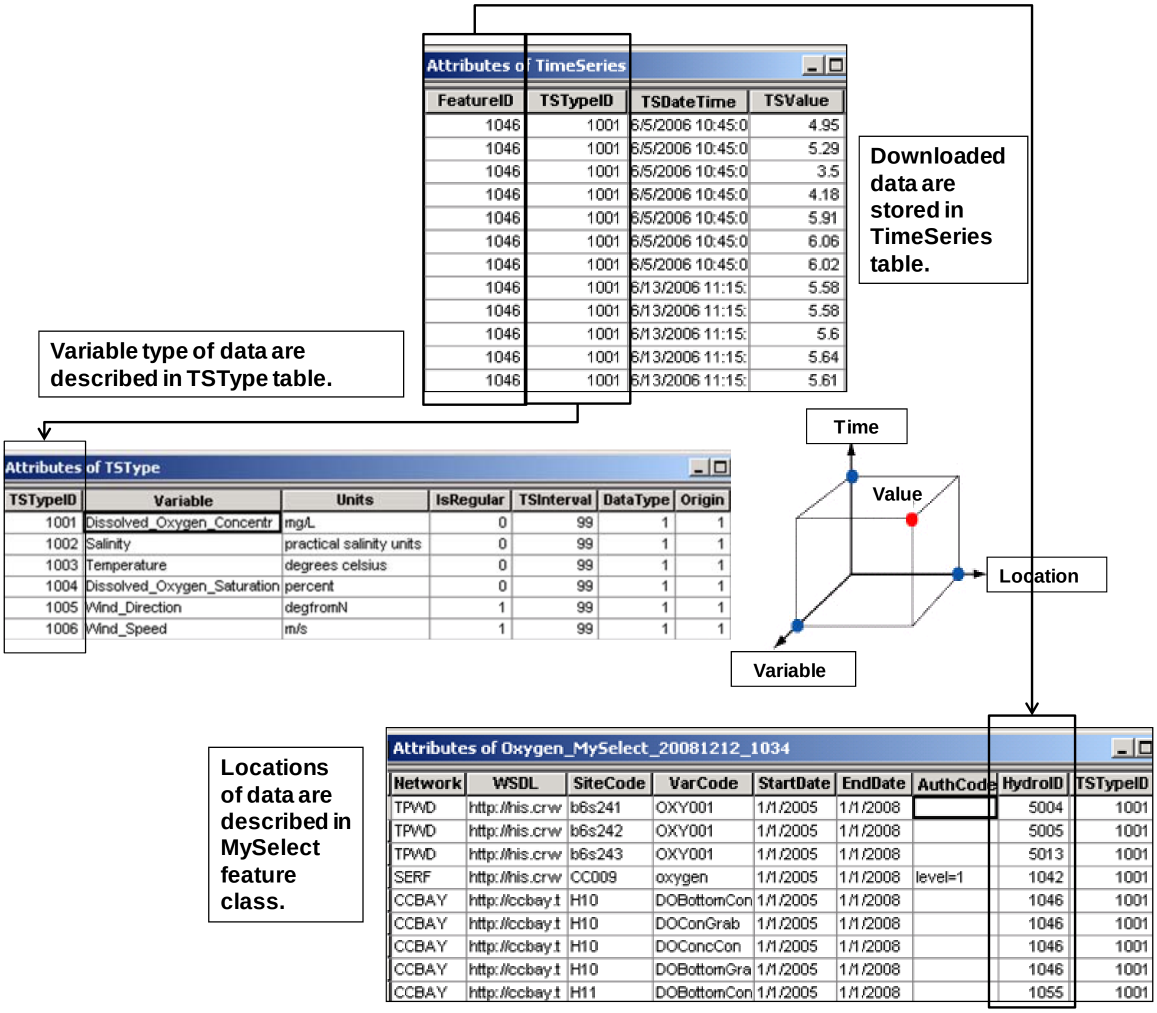
The feature class is converted to a MySelect feature class by appending extra fields (recall Tables 2.1 and 2.2) to its attribute table so that it was compliant with the MySelect table schema. A portion of the resulting table is shown in Figure 2.9. Note that because this particular MySelect table harvested dissolved oxygen data only, one TSTypeID (1001) is used to identify the variable type of all the downloaded data.

Attributes of Oxygen_MySelect_20081212_1034

Network	WSDL	SiteCode	VarCode	StartDate	EndDate	AuthCode	HydroID	TSTypeID
TPWD	http://his.crw	b6s241	OXY001	1/1/2005	1/1/2008		5004	1001
TPWD	http://his.crw	b6s242	OXY001	1/1/2005	1/1/2008		5005	1001
TPWD	http://his.crw	b6s243	OXY001	1/1/2005	1/1/2008		5013	1001
SERF	http://his.crw	CC009	oxygen	1/1/2005	1/1/2008	level=1	1042	1001
CCBAY	http://ccbay.t	H10	DOBottomCon	1/1/2005	1/1/2008		1046	1001
CCBAY	http://ccbay.t	H10	DOConGrab	1/1/2005	1/1/2008		1046	1001
CCBAY	http://ccbay.t	H10	DOConcCon	1/1/2005	1/1/2008		1046	1001
CCBAY	http://ccbay.t	H10	DOBottomGra	1/1/2005	1/1/2008		1046	1001
CCBAY	http://ccbay.t	H11	DOBottomCon	1/1/2005	1/1/2008		1055	1001

Figure 2.9 MySelect table that harvests oxygen data in southeast Corpus Christi Bay.

HydroGET is used to process the MySelect table and store the data into an Arc Hydro geodatabase. Downloaded time series are stored into the TimeSeries table. A new variable (i.e. TSTypeID: 1001, Dissolved Oxygen Concentration) is added to the TSType table. The MySelect, TSType, and TimeSeries tables together describe the location, variable, time and value of each dissolved oxygen observation (see Figure 2.10).



Attributes of TimeSeries

FeatureID	TSTypeID	TSDateTime	TSValue
1046	1001	6/5/2006 10:45:0	4.95
1046	1001	6/5/2006 10:45:0	5.29
1046	1001	6/5/2006 10:45:0	3.5
1046	1001	6/5/2006 10:45:0	4.18
1046	1001	6/5/2006 10:45:0	5.91
1046	1001	6/5/2006 10:45:0	6.06
1046	1001	6/5/2006 10:45:0	6.02
1046	1001	6/13/2006 11:15:	5.58
1046	1001	6/13/2006 11:15:	5.58
1046	1001	6/13/2006 11:15:	5.6
1046	1001	6/13/2006 11:15:	5.64
1046	1001	6/13/2006 11:15:	5.61

Attributes of TSType

TSTypeID	Variable	Units	IsRegular	TSInterval	DataType	Origin
1001	Dissolved_Oxygen_Concentr	mg/L	0	99	1	1
1002	Salinity	practical salinity units	0	99	1	1
1003	Temperature	degrees celsius	0	99	1	1
1004	Dissolved_Oxygen_Saturation	percent	0	99	1	1
1005	Wind_Direction	degfromN	1	99	1	1
1006	Wind_Speed	m/s	1	99	1	1

Attributes of Oxygen_MySelect_20081212_1034

Network	WSDL	SiteCode	VarCode	StartDate	EndDate	AuthCode	HydroID	TSTypeID
TPWD	http://his.crw	b6s241	OXY001	1/1/2005	1/1/2008		5004	1001
TPWD	http://his.crw	b6s242	OXY001	1/1/2005	1/1/2008		5005	1001
TPWD	http://his.crw	b6s243	OXY001	1/1/2005	1/1/2008		5013	1001
SERF	http://his.crw	CC009	oxygen	1/1/2005	1/1/2008	level=1	1042	1001
CCBAY	http://ccbay.t	H10	DOBottomCon	1/1/2005	1/1/2008		1046	1001
CCBAY	http://ccbay.t	H10	DOConGrab	1/1/2005	1/1/2008		1046	1001
CCBAY	http://ccbay.t	H10	DOConcCon	1/1/2005	1/1/2008		1046	1001
CCBAY	http://ccbay.t	H10	DOBottomGra	1/1/2005	1/1/2008		1046	1001
CCBAY	http://ccbay.t	H11	DOBottomCon	1/1/2005	1/1/2008		1055	1001

Figure 2.10 Dissolved oxygen in an ArcHydro database.

To plot the downloaded dissolved oxygen data, the Datacube diagram is devised. A data cube diagram is a set of graphs that plot the observed values along four dimensions – x (longitude), y (latitude), z (depth) and t (time). The four graphs are arranged around a map of the sampling area, which is essentially a graph of latitude vs. longitude. The axes of the "Value vs. Longitude" and "Latitude vs. Value" graphs are aligned with the latitude and longitude axes of the map. This allows data on the plots to be referenced directly to map features, namely the sensor locations. An example of a data cube diagram is shown in Figure 2.11. It shows all the dissolved oxygen data collected in southeast Corpus Christi Bay in August, 2005. The top graph shows the time series profile. The "strings" of data indicate days of collection on 8/3/2005, 8/9/2005, 8/16/2005, 8/23/2005 and 8/31/2005. The left and bottom graphs show profiles along the latitudinal and longitudinal axes. The "strings" of data in the left and bottom graphs can be extended into the map to pin point the location where they were collected. The right graph shows the depth profile of the oxygen. Oxygen values are observed to decrease as depth increased. Some of the oxygen values were less than 2 mg/L, indicating the occurrence of hypoxia during the period.

The space-time data cube structure of the Arc Hydro data model is very amenable to the display of spatial and temporal patterns of the data using the Datacube diagram. The Datacube diagram shown in Figure 2.11 is developed in Excel and retrieves data from an Arc Hydro geodatabase using SQL queries.

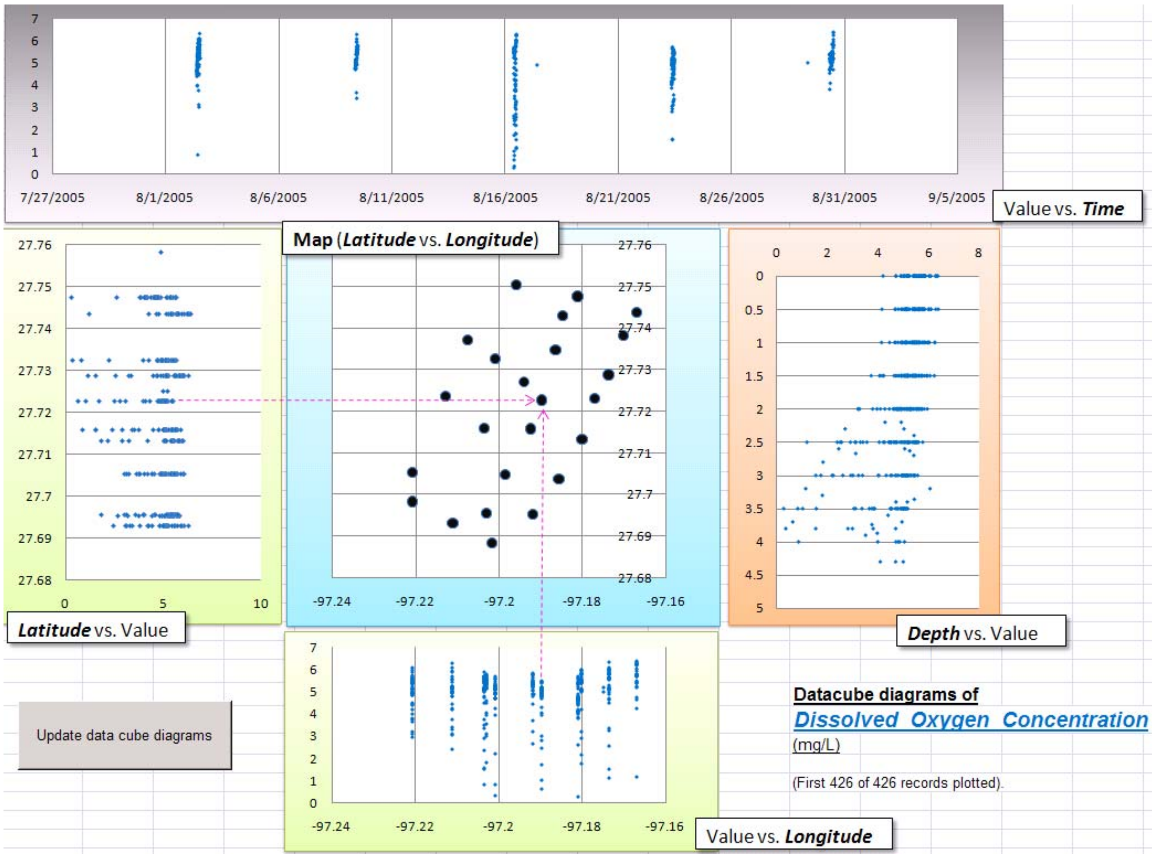


Figure 2.11. A data cube diagram of dissolved oxygen collected in southeast Corpus Christi Bay in August, 2005.

2.6 Conclusions for Chapter 2

This study describes how challenges in consolidating environmental data for hypoxia research in Corpus Christi Bay are met by the implementation of CUAHSI's Hydrologic Information System (HIS) and the HydroGET tool.

HIS opens up the data sources in the Bay to the researchers by unifying the methods for accessing their data. In the implementation of HIS in the Bay, two kinds of CUAHSI web services are used. Generic OD webservices are simple to implement but cannot handle continuous updates to the data source. Therefore they work better with data sources that are static or infrequently updated. Hybrid web services are more difficult to implement because they require custom programming of webscrapers. However, they can handle data sources that are constantly updated by real-time data.

HydroGET consolidates the data from the sources into a structured database to support scientific research. To store the downloaded data, HydroGET uses the Arc Hydro data model. The time series component of the Arc Hydro data model is versatile enough to accommodate any time series data that are collected at point locations. In addition, its structure is also amenable to the display of spatial and temporal data using the Datacube diagram. To manage the multiple web service requests involved in data harvesting, HydroGET employs the MySelect table. The MySelect table stores the parameters necessary for querying CUAHSI data web services. The table also works as a lookup table that mediates semantic difference between different data sources.

CHAPTER 3: USING SPACE-TIME INTERPOLATION TO CHARACTERIZE MOVEMENT OF GRAVITY CURRENTS IN CORPUS CHRISTI BAY

By Sin Chit To and David R. Maidment

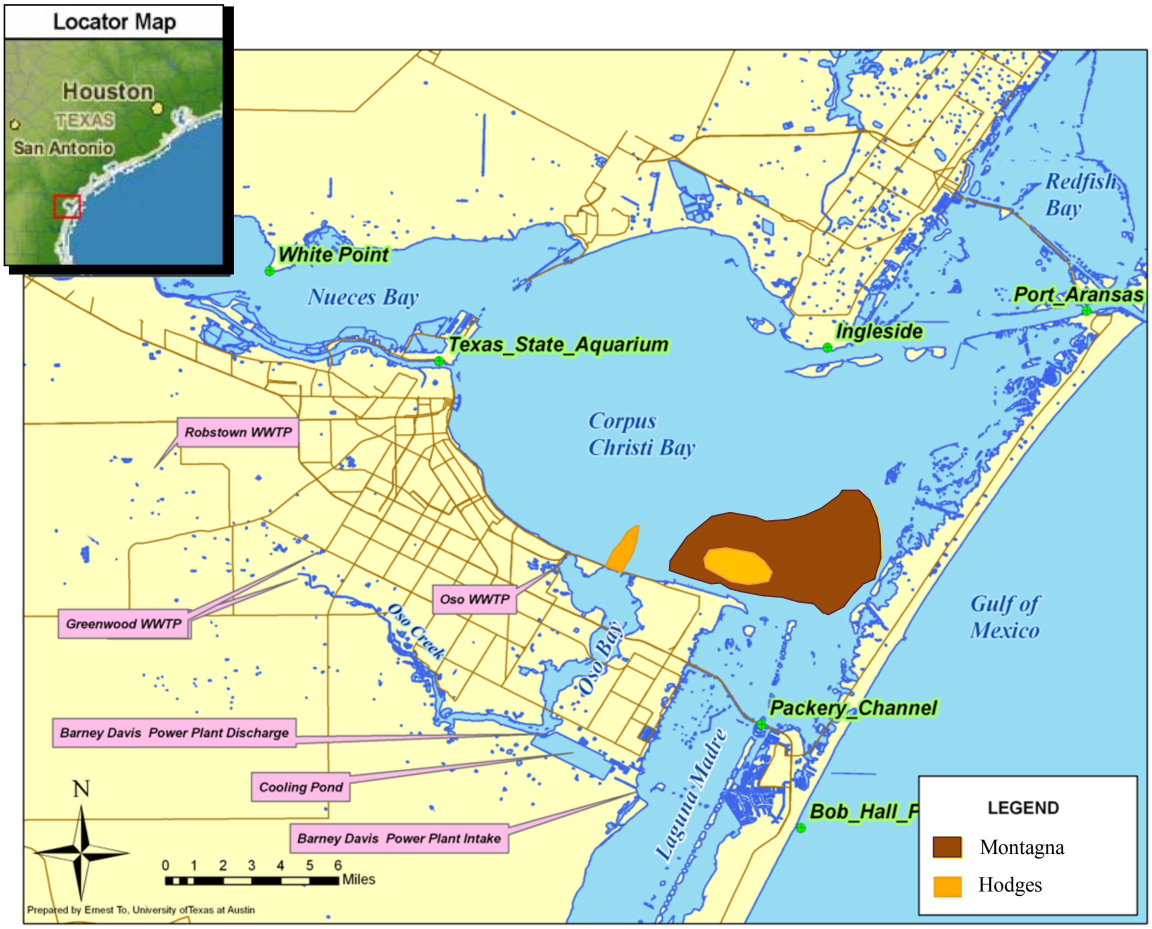
3.1 Abstract

In Corpus Christi Bay in south Texas, hypoxia is correlated with salinity-induced density stratification of the water column. One of the suspected causes of stratification is the introduction of gravity currents from shallow bays adjacent to Corpus Christi Bay. This paper describes the application of space-time interpolation models, namely kriging, for synthesis. The purpose of synthesis is to facilitate the construction of a continuous space-time volume of salinity. By dissecting the domain in regular time intervals, snapshots of the gravity current as it undergoes stages of emergence, movement and dissipation are visualized. From the snapshots, the persistence of a gravity current in the Bay is estimated to be on the order of 5 days (approximately one week). The gravity current speed is estimated to be on the order of magnitude of 1 km/day.

3.2 Introduction

MOTIVATION

Gravity currents have been observed in Corpus Christi Bay in south Texas (David and Hodges, 2006, Brower, et al., 2007, Hodges, et al., 2008) and are suspected to cause hypoxia in the Bay. In fluid dynamics, a gravity current is a primarily horizontal flow in a gravitational field that is driven by a horizontal density difference (Simpson, 1999). Density differences in natural water bodies can be caused by factors such as differences in temperature, salinity and sediment content. Gravity currents can have significant environmental impact on estuaries, lakes and oceans because they can cause vertical stratification in water bodies, which limits the transport of chemicals, such as dissolved oxygen, along the depth of the water column. For instance hypoxia can occur because dissolved oxygen near the surface of the water cannot reach the lower depths. Also, gravity currents can transport chemicals and physical properties (such as heat) horizontally over large horizontal distances. For example, turbidity currents on the seafloor can carry material thousands of miles. Gravity currents in Corpus Christi Bay consist of hypersaline water that flow into Corpus Christi Bay from adjacent water bodies, namely, Laguna Madre and Oso Bay (see Figure 3.1).



Ernest Sin Chit To, CRWR

Figure 3.1. Hypoxia in Corpus Christi Bay

Physically, Corpus Christi Bay is approximately circular in shape with a diameter of approximately 13 miles. The average depth of the Bay is 9 feet, and the Bay is separated from the Gulf of Mexico by a barrier island, so that water circulation in the Bay is driven more by wind than by tides.

In Figure 3.1, the yellow and orange areas in the southeast corner of the Bay indicate areas where episodic hypoxia has been found in the past. These areas are downstream of

Laguna Madre and Oso Bay. These are also areas where gravity currents have been found. Laguna Madre is a lagoonal system that is characterized by:

1. Shallow depth. The northern portion of it is about 3 to 4 feet deep – therefore shallower than Corpus Christi Bay.
2. Limited freshwater inflow. Most inflows are from surrounding surface runoff during storm events
3. Dense aquatic vegetation (Montagna, 1993).

Due to the shallowness of Laguna Madre and Corpus Christi Bay, as well as their limited interaction with the Gulf of Mexico, elevated salinity levels are found during periods in the summer when evaporation is high and precipitation is low. The typical salinity in the Gulf of Mexico ranges from 10 to 30 psu (practical salinity units). During the summer, salinity in Corpus Christi Bay can reach as high as 50 psu, while salinity in upper Laguna Madre can reach as high as 70 psu. The difference in salinity between the two bays results in a horizontal density difference that leads to the development of gravity currents. Oso Bay is an estuarine system that receives freshwater from Oso Creek. However it receives a flow of 500 MGD of hypersaline water from Laguna Madre through the Barney Davis power plant. The Barney Davis power plant draws cooling water from Laguna Madre and discharges it into a cooling pond that leads to Oso Creek. As a result, Oso Creek also experiences hypersalinity during the summer.

Role of gravity currents in hypoxia

Hypoxia has been found to be correlated with the occurrence of salinity-induced vertical stratification of the water column (Russell and Montagna, 2007). However, the exact mechanism of how gravity currents lead to hypoxia in the Bay has yet to be found. The following is a hypothesis of how gravity currents cause hypoxia (see Figure 3.2):

Dense aquatic vegetation in Laguna Madre retards the movement of hypersaline waters into Corpus Christi Bay. As a result, external forces, such as wind and tide, are needed to overcome the resistance. Once gravity currents are released, they travel down slope towards deeper areas of the Bay (see Figure 3.2). Transfer of dissolved oxygen into gravity currents is limited because of the energy barrier posed by the difference in densities between the hypersaline water in the gravity current and the less saline water above it. A net reduction in oxygen within the gravity current occurs because oxygen is depleted much faster by sediment oxygen demand than it is replenished by diffusion from the overlying water. If left unabated, oxygen concentration in the gravity current can drop to hypoxic levels. Despite the role of the wind as an initiator of gravity currents, it can also break up gravity currents by introducing mixing energy into the water column. As a result it can also prevent or stop hypoxia.

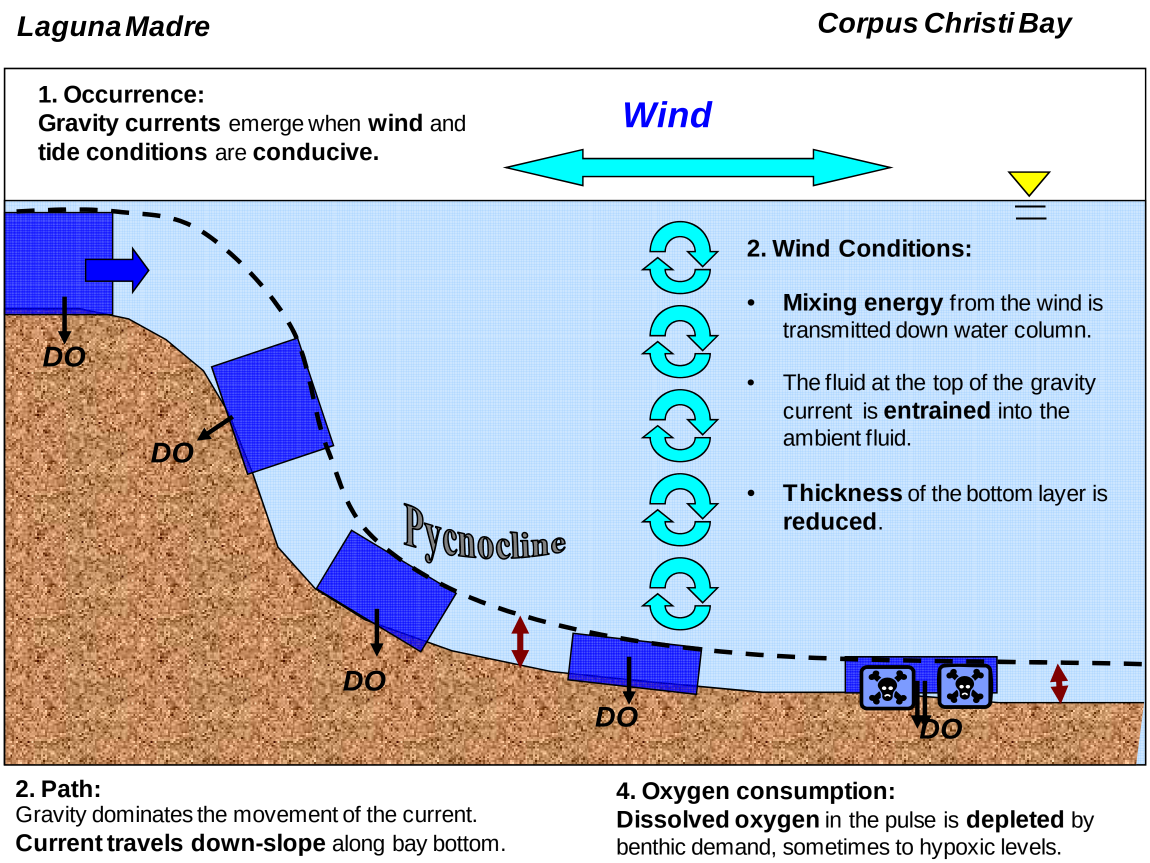


Figure 3.2. Relationship between gravity currents and hypoxia in southeast Corpus Christi Bay

In summary, four conditions need to be satisfied in order for hypoxia to happen at a given location in the Bay. These are *initiation, travel, wind conditions,* and *oxygen consumption.* First of all, a gravity current needs to emerge from the shallow bays. Secondly, the given location must be within the travel path of the current. Thirdly, wind conditions are not strong enough to break up the gravity current before it reaches the location. Lastly, dissolved oxygen is depleted below 2 mg/L when the gravity current reaches the location. An illustration is provided in Figure 3.3.

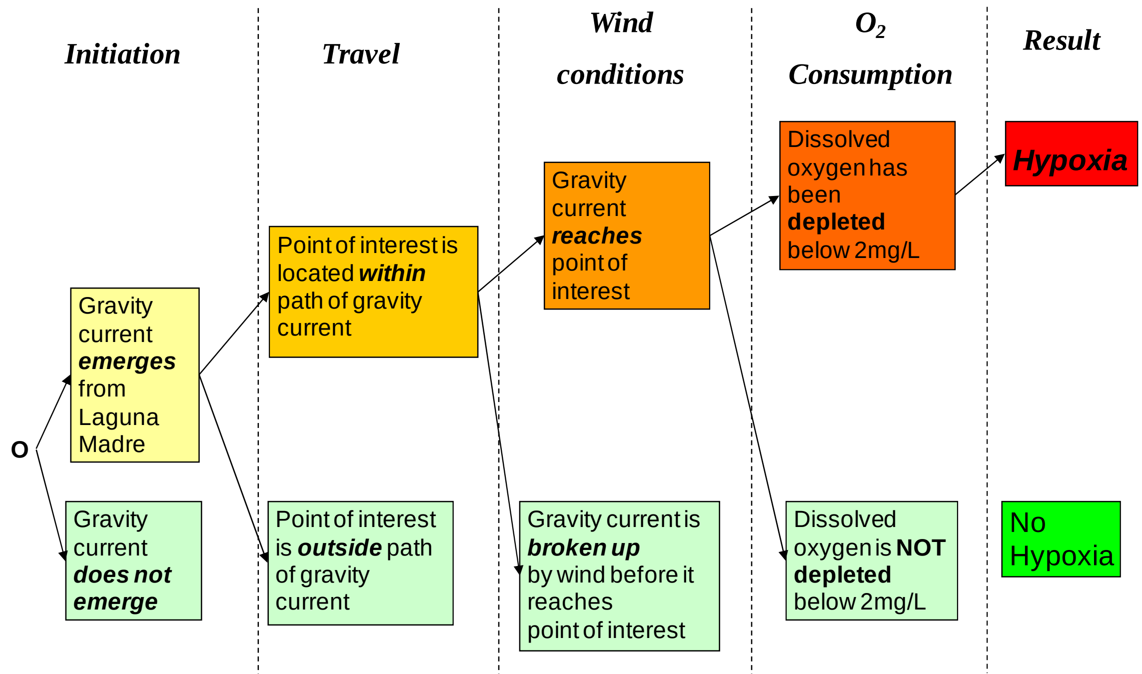


Figure 3.3. Conditions leading to hypoxia

MOTIVATION

Characterization of gravity currents

Since the 1970s, several governmental agencies and academic entities have deployed sensors in the Bay to monitor salinity and oxygen. Each sensor network measured a different region of the Bay at different spatial and temporal resolutions. This resulted in a collection of data sets – each set captured an incomplete piece of the hypoxia phenomenon. Despite this, no documented effort has been made in integrating these sets of data. Therefore the first goal of this interpolation study was to synthesize these fragments over space and time into a continuous description of the salinity in the Bay. By visualizing the synthesized results, physical properties of the gravity current, such as its speed and persistence through time, can be observed.

Evaluation of space-time kriging as a method for synthesizing data harvested using cyberinfrastructures

In recent years, efforts have been made to develop cyberinfrastructures to streamline and unite access to environmental data from multiple governmental and research organizations. CUAHSI's (Consortium of Universities for the Advancement of Hydrologic Science, Inc) Hydrologic Information System (HIS) is one of these cyberinfrastructures. HIS is a system that provides researchers with a unified set of tools and protocols to publish and access data over the internet. The number of organizations around the nation that have adopted HIS has steadily increased since HIS's inception in 2007. Within Corpus Christi Bay alone, data from 5 sensor networks have been published via HIS. The need for synthesizing data from multiple sources is growing and the challenge is to find an appropriate method. Kriging is a widely used geostatistical method for interpolating spatial data and is very popular in oil and mining applications. Its strength lies in the fact that it is a statistical method and therefore general enough to be applied to a broad range of variables. In southeast Corpus Christi Bay, it has been previously used to interpolate salinity data from one of the sensor networks in three spatial dimensions (To, 2007). However, interpolation across the time dimension has not yet been attempted for the Bay. Space-time kriging is a relatively recent development in geostatistics. Its theoretical foundations were expounded by mathematicians such as Rouhani and Meyers (1990), Cressie and Huang (1999) and De Iaco, et al., (2000, 2002). Its application has been made possible via tools developed by researchers like Christakos, *et al.* (2002) and DeCesare, *et al.* (2001). Because of its recent development, its application to natural environmental processes, particularly those in estuarine systems, is still limited. The second goal of this study is therefore to evaluate space-time kriging as a method for synthesizing data harvested using cyberinfrastructures.

3.3 Scope of investigation

STUDY AREA

The area selected for the interpolation study is situated in southeast Corpus Christi Bay (see area outlined by dash line in Figure 3.4). In 2007, NOAA published new bathymetry data in Corpus Christi Bay as part of its SIFT (Short-Term Inundation Forecasting for Tsunami) project (NOAA, 2007b). The bathymetry data indicated that the bottom of the Bay is uneven and is populated with underwater ridges and depressions. These terrain features can be used to compartmentalize the Bay into multiple underwater basins using the same principles as watershed delineation. In southeast Corpus Christi Bay, one such underwater basin which receives hypersaline waters from the eastern portion of Laguna

Madre is identified. As a result it presents an isolated system where one may observe gravity currents emerging from one shallow bay (i.e. East Laguna Madre). For the rest of this paper, this region referred to as the study area.

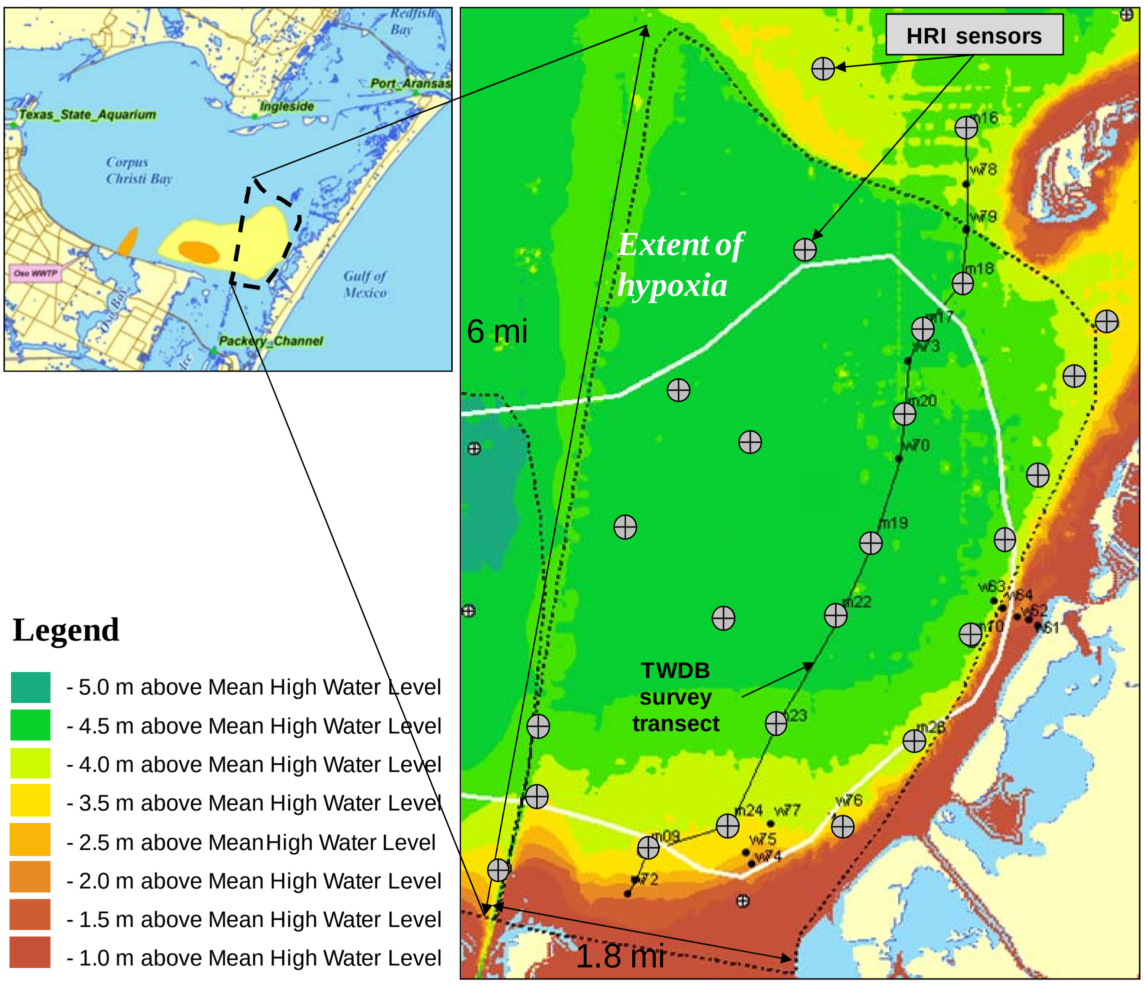


Figure 3.4. The study area

SOURCES OF SALINITY DATA

Within the study area are two sensor networks that measure salinity. These are the HRI network and TWDB network. The HRI network is operated by Dr. Paul Montagna of Texas A&M Corpus Christi. It contains data collected by Dr. Paul Montagna since the 1980s. The TWDB network is operated by Dr. Ben Hodges of University of Texas under the funding of the Texas Water Development Board. Table 3.1 provides a brief description of each network and the data they collect.

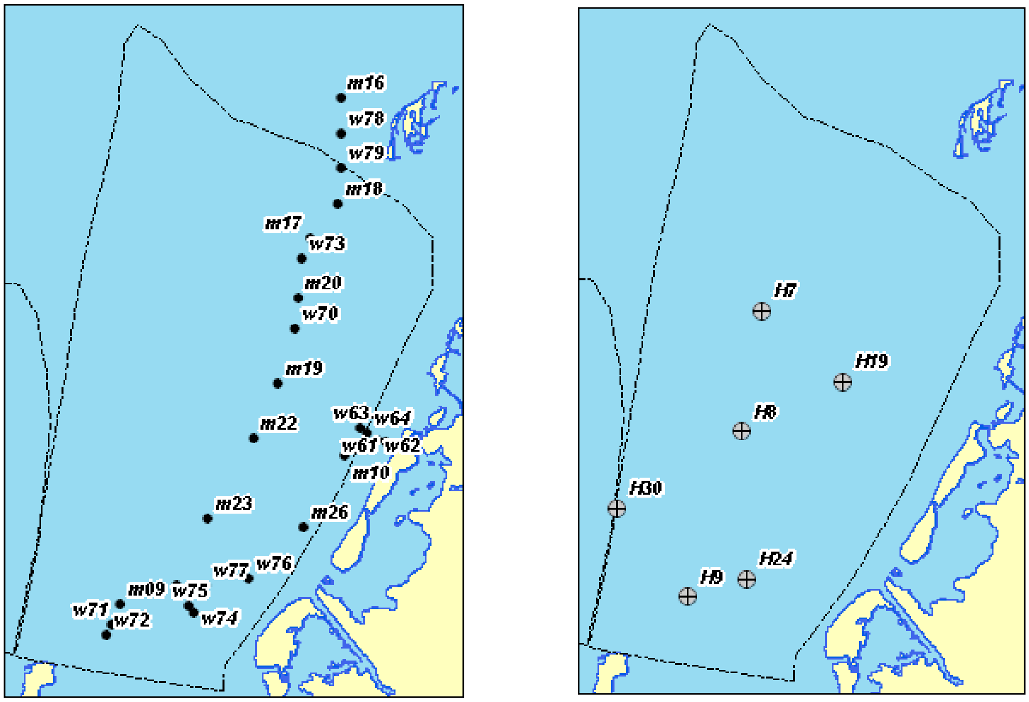
Table 3.1. Sensor networks in study area of Corpus Christi Bay

	Network	Description	Data	Locations within study area	Sampling period	Sampling frequency
1	HRI	• Harte Research Institute • Operated by Dr. Paul Montagna of Texas A&M, Corpus Christi • Previously known as the UTMSI network – also operated by Dr. Montagna.	Depth profiles of • salinity • dissolved oxygen • temperature • other water quality data.	23 fixed stations	Every summer since the 1980s.	Bi-weekly to monthly
2	TWDB	• Plume tracking studies • Operated by Dr. Ben Hodges, University of Texas at Austin	Depth profiles of • salinity • dissolved oxygen • temperature	Multiple locations along transect	July 14 to 17, 2006.	Twice daily

STUDY PERIOD

Between the two sensor networks, HRI possesses data for the longer period of time and over the largest area. With its sampling frequency of roughly two weeks it has been able to detect the presence of salinity stratification in the Bay. However, the resolution is not fine enough to track the movement and development of gravity currents. On the other hand, Dr. Hodges performed one plume tracking survey in the study area from July 14 to

17, 2006. Not only did this survey detect a gravity current, the survey's sampling frequency of twice daily meant that it was able to capture the movement of the gravity current. For this reason, the study period for this interpolation was defined around the period July 14 to 17, 2006. Serendipitously, HRI collected samples on July 12, and 18, 2006 in the study area. These data provide more information on the initiation and dissipation of the gravity current. By combining these two data sets and then synthesizing them using space-time interpolation, the history of the gravity current as it underwent initiation, movement and dissipation can be observed. Figure 3.5 shows the location of the TWDB survey transect and the HRI stations that were sampled during this period.



Plume tracking survey July 14 to 17, 2006. (During movement of gravity current)

Water quality data July 12 and 18, 2006. (At initiation and demise of gravity current)

Figure 3.5. Sampled locations in the southeast Corpus Christi Bay study area from July 12 to 18, 2006 (To, 2008)

3.4 Data Preparation

DATA HARVESTING

Both HRI and TWDB data are published via web services of the CUAHSI hydrologic information system. A web service client called HydroGET (To, 2008) harvested data from these two sources into a Microsoft Access database that followed the Arc Hydro data model (Maidment, 2002) structure. An illustration of the data harvesting process is shown in Figure 3.6.

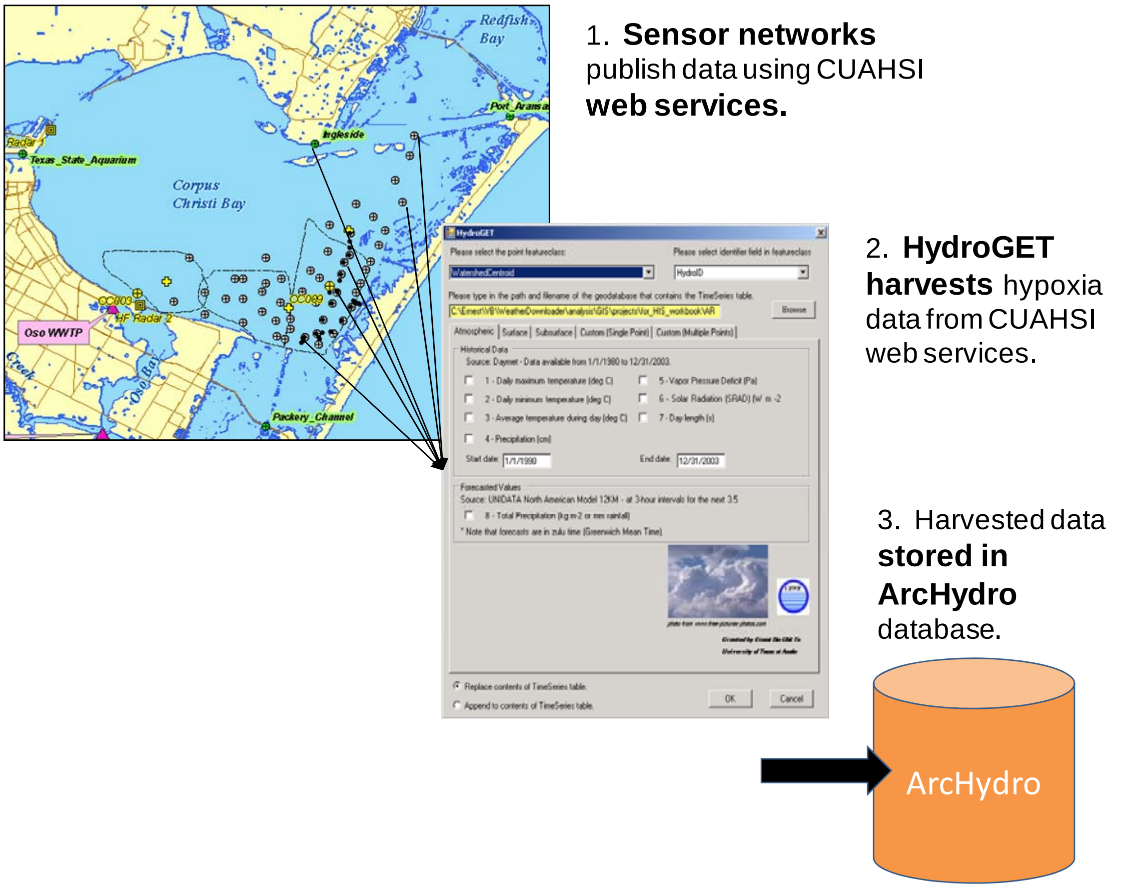


Figure 3.6. Harvesting sensor data into a database.

PROJECTION OF DATA ONTO DIRECTION OF FLOW

Because of the way the data were collected, data locations are distributed densely along the TWDB transect but sparsely in the transverse direction. Due to this non-uniformity, preliminary attempts at interpolating across the entire surface area yielded anomalous results. Therefore the original intention of interpolation along all four dimensions (i.e., latitude, longitude, depth and time) was not pursued. To improve the results, the following simplifying assumption was made to reduce the number of dimensions from four to three:

Recall from Figure 3.4, that the bottom of the study area sloped gently downwards along a north to north-easterly direction. This suggested that gravity currents flowed along this general direction and that the front of the current was oriented in a direction approximately perpendicular to the flow.

By assuming that the depth profile of salinity was relatively uniform along the front, the number of spatial dimensions was reduced from three to two by omitting the transverse direction. This was achieved by defining a reference axis oriented along the flow direction (see Figure 3.7). Data locations were perpendicularly projected onto the axis and then assigned a distance value which corresponded to the length between the projected location and the origin of the axis. The origin of the axis was set at the mouth of East Laguna Madre. Thus the original four dimensions (latitude, longitude, depth and time) are reduced to just three (along the direction of flow, depth and time).

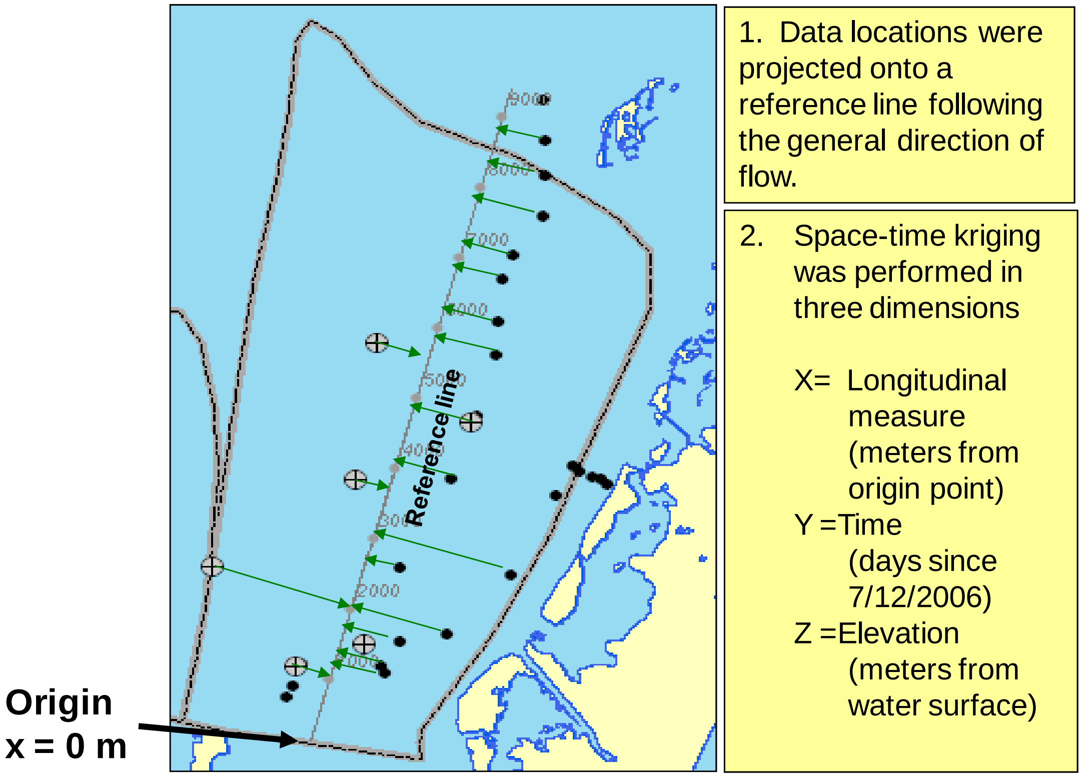


Figure 3.7. Synthesis of salinity data collected in East SECCB Study Area.

3.5 Methodology

BACKGROUND ON SPACE-TIME KRIGING

Space-time kriging

Space-time kriging extends conventional kriging into the time dimension. It is fundamentally the same as conventional spatial kriging except in the management of covariance modeling (De Cesare, *et al*., 2001, Christakos, *et al*., 2002). One may seek to treat time as if it is an extra dimension in spatial kriging and then lump the lag distances along the spatial and temporal axes into a single Euclidean space-time metric. However, this has been shown to be theoretically and practically problematic (Rouhani and Myers, 1990). The key reason is because the time axis is by nature different (i.e. it is not geometric) and not necessarily orthogonal to the other three principal axes (Srinivasan, 2007). Environmental processes occurring in time almost always have some dependence on processes occurring in space, which accounts for some of the non-orthogonality between the time and spatial axes.

Covariance modeling in space-time kriging

When orthogonality and similarity among the axes of the random space are in question, it is necessary to step back to a more general form and treat the lags along the different axes as separate arguments. As a demonstration, let us revisit three dimensional spatial kriging and start from the general form of the covariance function in three-dimensional space, which is:

$$C(u_1, u_2) = f((x_1 - x_2), (y_1 - y_2), (z_1 - z_2)) \qquad \text{Equation 3.1}$$

Where:

u_1 and u_2 are the locations of any given two data points, 1 and 2.

x_1, x_2, y_1, y_2, z_1, z_2 are the coordinates of points 1 and 2 along the three geometric axes.

When the covariance function is directionally independent (i.e. isotropic), the lags along the axes can be summarized into a single lag, parameter, h, using the distance formula, where

$$h = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \qquad \text{Equation 3.2}$$

When the covariance function is directionally dependent (i.e. anisotropic), linear transformations are used to project the lag vector – which is the vector between two locations – onto a space where the three principal axes are 1) along the direction of the greatest correlation and 2) minor direction of correlation and 3) the direction normal to the plane defined by the former two directions. To account for different correlation distances, a subsequent linear transformation is applied to scale the transformed vector to isotropic space. The mathematical functions that perform these transformations are described in many geostatistics text books and are not presented here because of their length. The transformed lag, h', can then be applied to Equation 3.1 to estimate the covariance. The resultant form of the covariance model is shown in Equation 3.3.

$$C(u_1, u_2) = f(h') \qquad \text{Equation 3.3}$$

Equation 3.3 is the common form of the covariance function used in spatial kriging where one spatial metric, h', is used to characterize lag between two points. Using the distance formula to calculate lag is only applicable when the x, y, z axes are orthogonal to each other. In space-time kriging, the temporal axis is not necessary orthogonal to the spatial axis because processes that happen in space often have dependencies on time. Therefore the appropriate method is to treat the spatial lag, h, and the temporal lag, τ, as independent arguments. As a result, the covariance function has the form:

$$C(u_1, u_2) = f(h', \tau) \qquad \text{Equation 3.4}$$

This treatment of the space-time covariance leads to two families of models. The first is the family of separable models where the covariance function can be separated into covariance functions that are either totally dependent on the spatial lag, h, or the temporal lag, τ . This family contains the product sum model shown in Equation 3.5 and the product model shown in Equation 3.6.

$$C(u_1, u_2) = f_1(h) f_2(\tau) + f_3(h) + f_4(\tau) \qquad \text{Equation 3.5}$$

$$C(u_1, u_2) = f_1(h) f_2(\tau) \qquad \text{Equation 3.6}$$

The second is the family of non-separable models that cannot be separated into covariance functions that are either totally dependent on the spatial lag, h, or the temporal lag, τ.

$$C(u_1, u_2) = f(h, \tau) \qquad \text{Equation 3.7}$$

According to De Iaco, et al., (2000, 2002) and Cressie and Huang (1999) both families of space-time covariance models are valid.

Tools for space-time kriging

The most common tool used in kriging is GSLIB which is developed by Clayton V. Deutch and Andre Journel from Stanford University (Deutch and Journel, 1992.). It is, however, limited to performing spatial kriging only. To execute space-time kriging, De Cesare*, et al.*(2002) augmented the original GSLIB libraries with extra FORTRAN 77 routines for performing space-time covariance modeling using the separable models shown in Equation 3.5. These tools include programs for plotting and fitting space-time variograms and programs for implementing separable space-time covariance models in kriging. This study uses De Cesare's version of GSLIB, which can be found at the website of the International Association of Mathematical Geology (De Cesare, et. al., 2002).

Unfortunately, De Cesare, *et al.* did not increase the total number of dimensions handled by GSLIB. The addition of the time dimension was compensated by the reduction of the number of spatial dimensions from three to two. Fortunately for this study, data points were already projected onto the direction of flow before the interpolation, thereby reducing the number of dimensions from four (northing, easting, depth and time) to three (distance from origin, depth and time). In this way, De Cesare's version of GSLIB was applicable to this study.

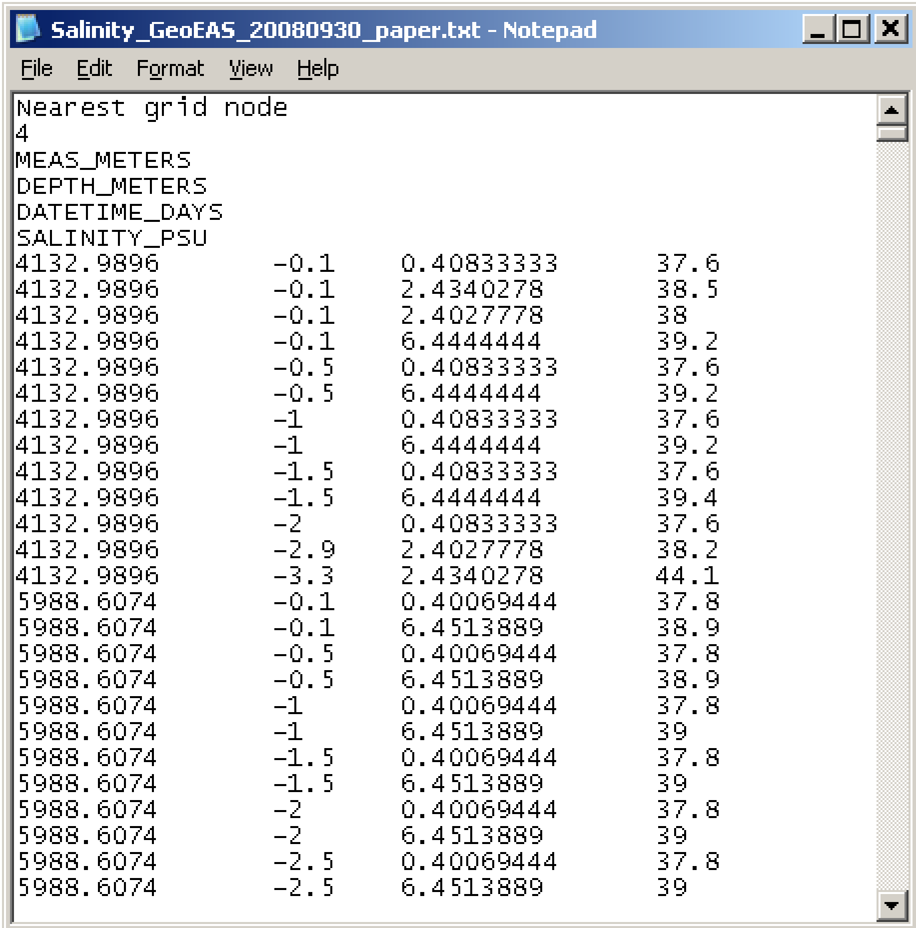
PROCEDURES

Because the GSLIB interface is very basic and has little means of operating with other computer programs, several scripts were written in ITTVIS's IDL programming language to manage the workflow of the overall geostatistical interpolation process. The workflow process consists of the following tasks:

1. Generating input data file from the Arc Hydro database
2. Write parameter files for the modified GSLIB functions, GAMVmod.exe and KT3Dmod.exe, to perform variogram analysis and space-time kriging.
3. Run the GAMVmod.exe and KT3Dmod.exe functions.
4. Visualize kriging results using voxels.

Data file

The combined salinity data were written GEOEAS format, which is the input data format for GSLIB. A snapshot of the data file is shown in Figure 3.8. The first column contains the x-coordinates of the samples, which are the distances (in meters) from the origin along the direction of flow. The second column contains the y-coordinates, which are the depths of the samples in meters. The third column contains the t-coordinates, which are the length of time (in days) since July 12, 2006 0000hrs (Central Standard Time). Finally, the fourth column contains the salinities measured (in practical salinity units). Altogether the table contains 1510 data points spanning 500 to 9300 m along the direction of flow, 0 to 4.3 m along the depth and 6.5 days. The average salinity is 40.8 psu and the sample variance is 3.6 psu^2.




```
Nearest grid node
4
MEAS_METERS
DEPTH_METERS
DATETIME_DAYS
SALINITY_PSU
4132.9896       -0.1    0.40833333      37.6
4132.9896       -0.1    2.4340278       38.5
4132.9896       -0.1    2.4027778       38
4132.9896       -0.1    6.4444444       39.2
4132.9896       -0.5    0.40833333      37.6
4132.9896       -0.5    6.4444444       39.2
4132.9896       -1      0.40833333      37.6
4132.9896       -1      6.4444444       39.2
4132.9896       -1.5    0.40833333      37.6
4132.9896       -1.5    6.4444444       39.4
4132.9896       -2      0.40833333      37.6
4132.9896       -2.9    2.4027778       38.2
4132.9896       -3.3    2.4340278       44.1
5988.6074       -0.1    0.40069444      37.8
5988.6074       -0.1    6.4513889       38.9
5988.6074       -0.5    0.40069444      37.8
5988.6074       -0.5    6.4513889       38.9
5988.6074       -1      0.40069444      37.8
5988.6074       -1      6.4513889       39
5988.6074       -1.5    0.40069444      37.8
5988.6074       -1.5    6.4513889       39
5988.6074       -2      0.40069444      37.8
5988.6074       -2      6.4513889       39
5988.6074       -2.5    0.40069444      37.8
5988.6074       -2.5    6.4513889       39
```

Figure 3.8. Snapshot of the input data file for space-time interpolaion

Variogram modeling

Three experimental variograms were plotted using the GAMVmod.exe function in DeCesare's modified GSLIB. The variograms are

1. variogram along the direction of flow;
2. variogram along the depth; and,

3. variogram along the time axis.

The experimental variograms were fitted by eye using professional judgment.

Figure 3.9 shows the fitted and experimental variograms along the direction of flow. The fitted variogram adopts the Gaussian model structure and its equation is:

$$\gamma_x(h_x) = \gamma_{0x} + (\gamma_{1x} - \gamma_{0x})(1 - e^{(\frac{-3h_x^2}{2a_x^2})}) \quad \text{Equation 3.8}$$

where

h_x = lag distance along direction of flow in meters

γ_{0x} = nugget = 2 psu^2

γ_{1x} = sill = 3.6 psu^2

a_x = range = 6000 m

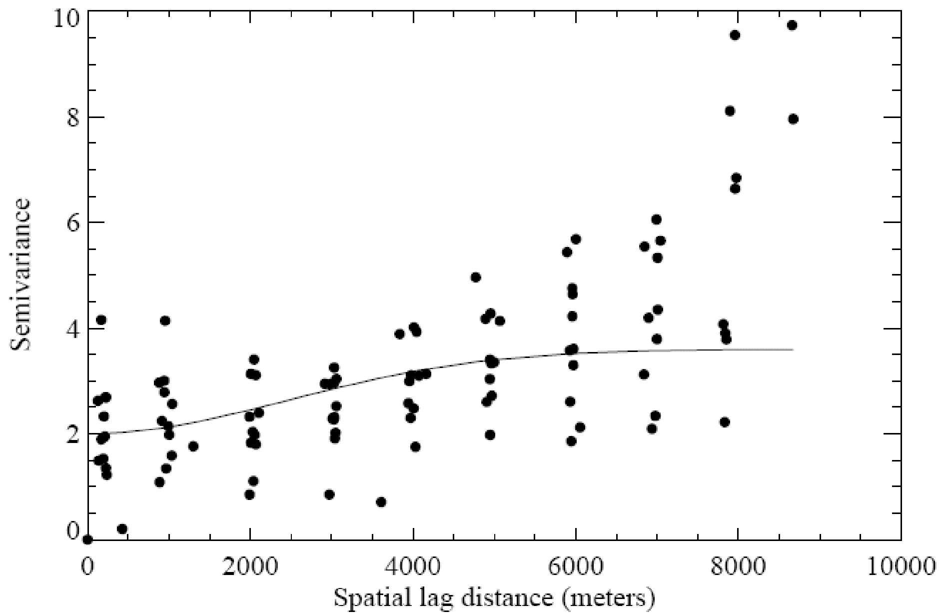


Figure 3.9. Experimental and fitted variograms along the longitudinal axis. (Spatial lag distances are in meters and semivariances are in psu^2).

Figure 3.10 shows the fitted and experimental variograms along the depth. The fitted variogram along the depth axis also adopts the Gaussian model structure and its equation is:

$$\gamma_y(h_y) = \gamma_{0y} + (\gamma_{1y} - \gamma_{0y})(1 - e^{(\frac{-3h_y^2}{2a_y^2})}) \qquad \text{Equation 3.9}$$

where

h_y = lag distance along the depth in meters

γ_{0y} = nugget = 0 psu^2

γ_{1y} = sill = 3.6 psu^2

a_y = range = 1.7 m

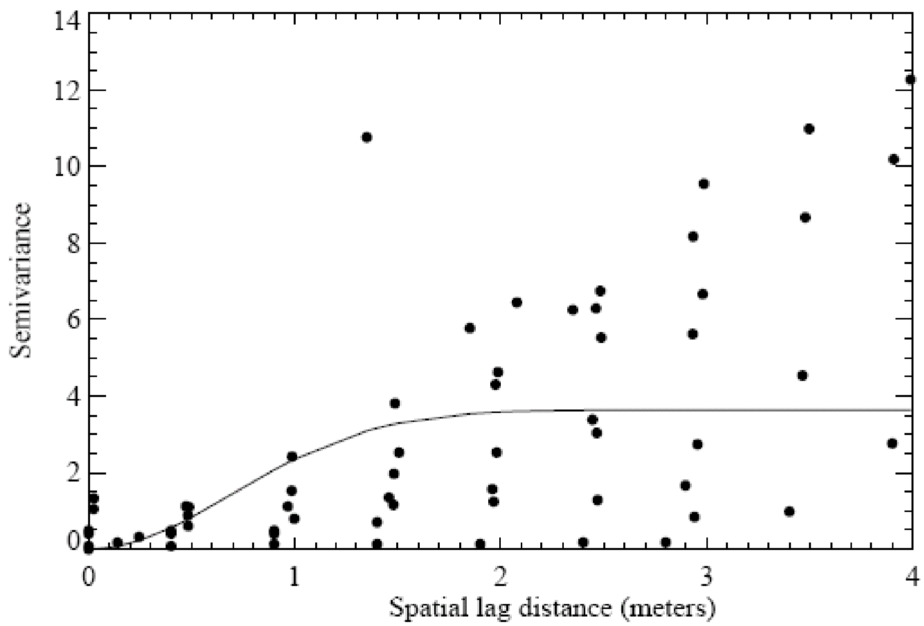


Figure 3.10. Experimental and fitted variograms along the depth axis. (Spatial lag distances are in meters and semivariances are in psu^2).

Figure 3.11 shows the fitted and experimental variograms along the time axis. The fitted variogram along the time axis uses the spherical model structure and its equation is:

$$\gamma_t(\tau) = \gamma_{0t} + (\gamma_{1t} - \gamma_{0t})(\frac{3\tau}{2a_t} - \frac{\tau^3}{2a_t^3})$$ Equation 3.10

where

τ = lag distance along the time axis (in days)

γ_{0t} = nugget = 0 psu^2

γ_{1t} = sill = 3 psu^2

a = range = 1 day

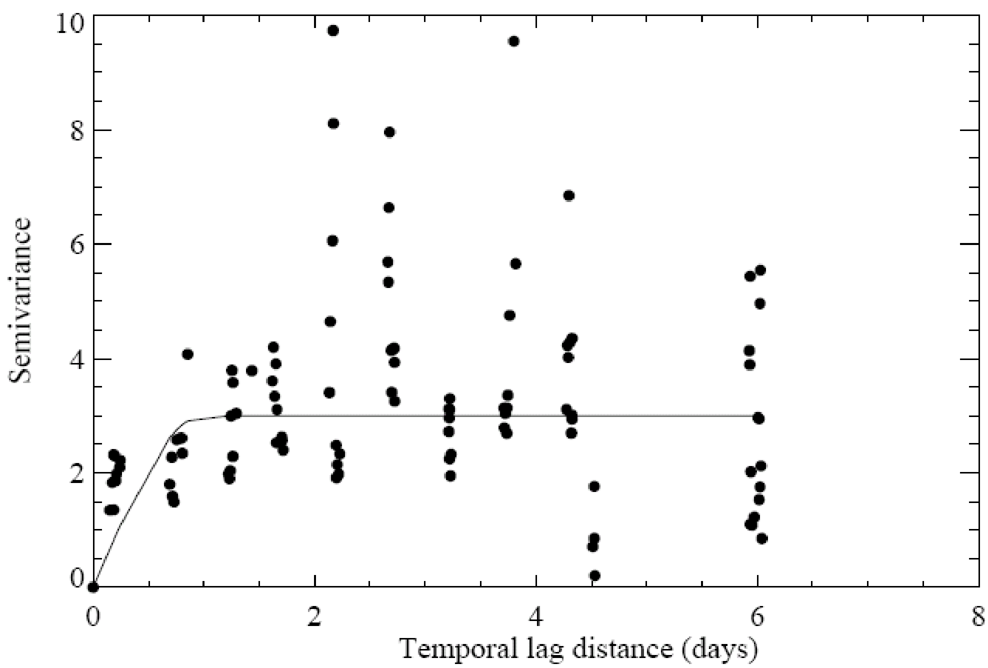


Figure 3.11. Experimental and fitted variograms along the time axis. (Temporal lag distance is in days and semivariances are in psu^2).

EXECUTION OF KRIGING

The modified GSLIB generates the spatial-temporal variogram model from the fitted variograms along the depth, direction of flow and temporal axes. First the variogram models for the two spatial axes are combined into a spatial variogram model of the form, $\gamma_s(h_x, h_y)$. Because the variogram along the direction of flow has a nugget, it is broken down into two structures. Structure 1 is for the nugget:

$$\gamma_{x1}(h_x, h_y) = 2(1 - e^{\frac{-3}{2}[(\frac{h_x}{\varepsilon})^2 + (\frac{h_y}{\infty})^2]})$$

Equation 3.11

Structure 2 is for the rise to the partial sill of 1.6 at a range of 6000 m.

$$\gamma_{x1}(h_x, h_y) = 1.6(1 - e^{\frac{-3}{2}[(\frac{h_x}{6000})^2 + (\frac{h_y}{\infty})^2]})$$

Equation 3.12

Where ε is an infinitesimally small positive number, that causes a discontinuity in the variogram (i.e. the nugget effect)

$\frac{h_y}{\infty}$ is a dummy term added so that the variogram model is nominally dependent on h_y.

For the variogram along the depth, the equation is:

$$\gamma_y(h_x, h_y) = 3.6(1 - e^{\frac{-3}{2}[(\frac{h_x}{\infty})^2 + (\frac{h_y}{1.7})^2]})$$

Equation 3.13

Summing Equations 3.11, 3.12 and 3.13 together yields the spatial variogram, $\gamma_s(\vec{h})$, (Equation 3.14):

$$\gamma_s(\vec{h}) = \gamma_s(h_x, h_y) =$$
$$2(1 - e^{\frac{-3}{2}[(\frac{h_x}{\varepsilon})^2 + (\frac{h_y}{\infty})^2]}) + 1.6(1 - e^{\frac{-3}{2}[(\frac{h_x}{6000})^2 + (\frac{h_y}{\infty})^2]})$$
$$+ 3.6(1 - e^{\frac{-3}{2}[(\frac{h_x}{\infty})^2 + (\frac{h_y}{1.7})^2]})$$

Equation 3.14

where $\vec{h}$ is a vector that consists of the elements, h_x and h_y .

The spatial variogram model, $\gamma_s(\vec{h})$, is combined with the temporal model , $\gamma_t(\tau)$, to produce the spatial-temporal model, $\gamma_s(h,\tau)$, using Equation 3.15 (De Cesare, et al, 2000):

$$\gamma_{st}(\vec{h},\tau) = k\gamma_{1t}\gamma_s(\vec{h}) + \gamma_{1s}\gamma_t(\tau) - \gamma_s(\vec{h})\gamma_t(\tau)$$

Equation 3.15

where

γ_{1t} = temporal sill = 3.6 psu

γ_{1t} = spatial sill = 3 psu

k is a parameter determined by Equation 3.16 (De Cesare, et al, 2000):

$$k = \frac{\gamma_{1st}}{\gamma_{1s}\gamma_{1s}}$$

Equation 3.16

where γ_{1st} is the global sill that can be estimated from sample data.

The variogram model in Equation 3.15 corresponds to the space-time covariance model, $C_{st}(\vec{h},\tau)$, shown in Equation 3.17.

$$C_{st}(\vec{h},\tau) = kC_s(\vec{h})C_t(\tau)$$ Equation 3.17

Where

$C_s(\vec{h})$ is the spatial covariance model

$C_t(\tau)$ is the temporal covariance model

Notice the form of Equation 3.17 corresponds with the product model shown in Equation 3.6.

This space-time covariance model in Equation 3.17 is then used in conjunction with the input data to produce the interpolation results.

3.6 Results and discussion

KRIGING RESULTS

A space-time volume comprising of three dimensions (distance, depth and time) was created after the execution of kriging (see Figure 3.12). By slicing out cross sections from the volume at different time steps, salinity profiles along the depth and direction of flow were obtained.

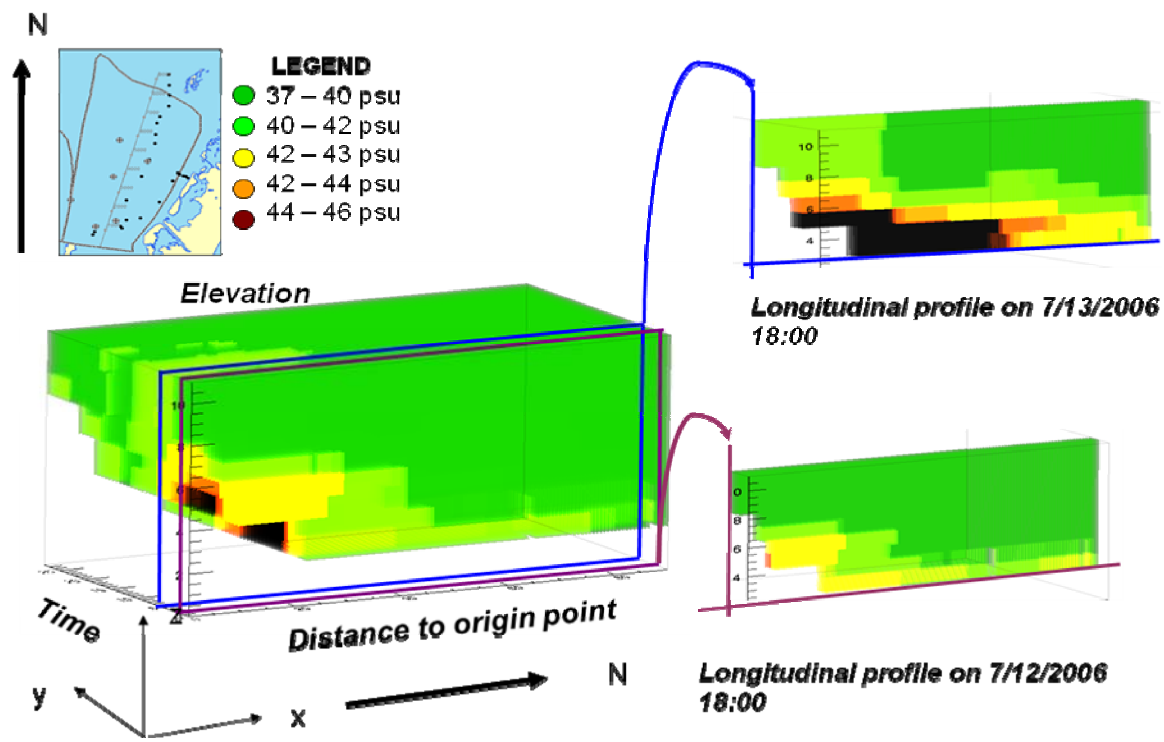


Figure 3.12. Workflow of computer programs for performing space-time kriging

Figure 3.13 shows a series of salinity profiles sliced at 12-hour intervals from the space-time volume. The black dots in each cross section depict the location of actual samples collected during the preceding time interval.

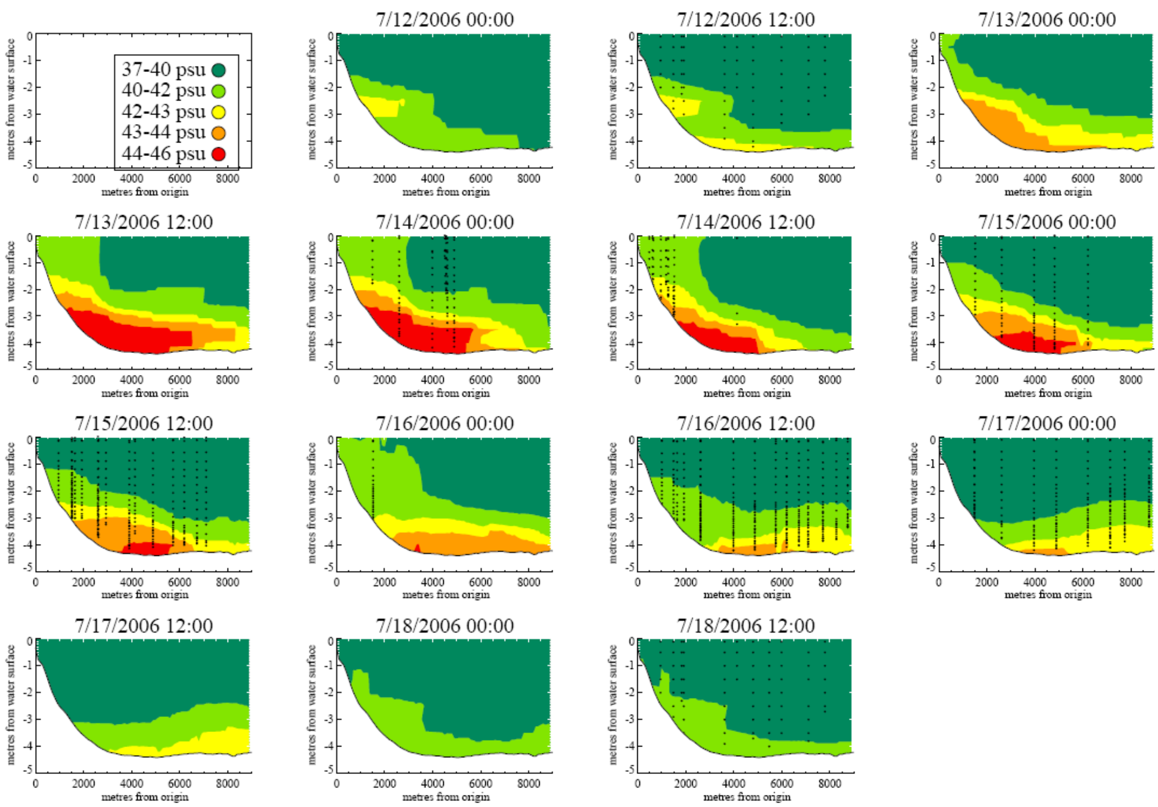


Figure 3.13. Snapshots of gravity current as it travels from the mouth of Laguna Madre (located at x-axis origin) into Corpus Christi Bay (To and Maidment, 2008).

The progress of the gravity current can be described as follows:

1. At the beginning in July 12, 2006, samples indicated slight salinity stratification in the study area. The highest salinity (42 to 43 psu) was found at mid-slope near the upstream

end and the lowest salinity (less than 40 psu) was found at the water surface at the downstream end.

2. On July 14, 2006, samples collected indicated the presence of high salinity water (44 to 46 psu) at the bay bottom between 2000 to 5000 m from the mouth of Laguna Madre. The region colored in red can be visually interpreted as the extent of the gravity current. Through space-time interpolation, the progress of the gravity current as it descended into the Bay between July 12 and 14 was visualized.

3. Between July 14 and July 16, the gravity current advanced further downstream. The salinity of the current was diluted during its movement due to lateral entrainment of less saline ambient fluid. For this period in time, the orange region (which represents the salinity range of 43 to 44 psu) can be visually interpreted as the extent of the gravity current. The gravity current was estimated to reach the downstream end of the study area on July 16.

4. Between July 16 and July 18, the gravity current was gradually dissipated. The thickness of the gravity current was slowly reduced to nothing. Samples collected on July 18 indicated slight stratification in the study area. The salinity profile on July 18 was similar to the state observed at the beginning of the study on July 12.

The time history of the gravity current can also be visualized from top down. By assuming that the front of the gravity current was perpendicular to the reference line, the propagation of the gravity current was mapped. Figure 3.14 shows the estimated bottom salinity in the Bay from at 12 hour intervals. The colored dots show the measured salinities at the bottom of the Bay. The series of profile shows the advancement and dissipation of the gravity current at the bottom of the Bay through time.

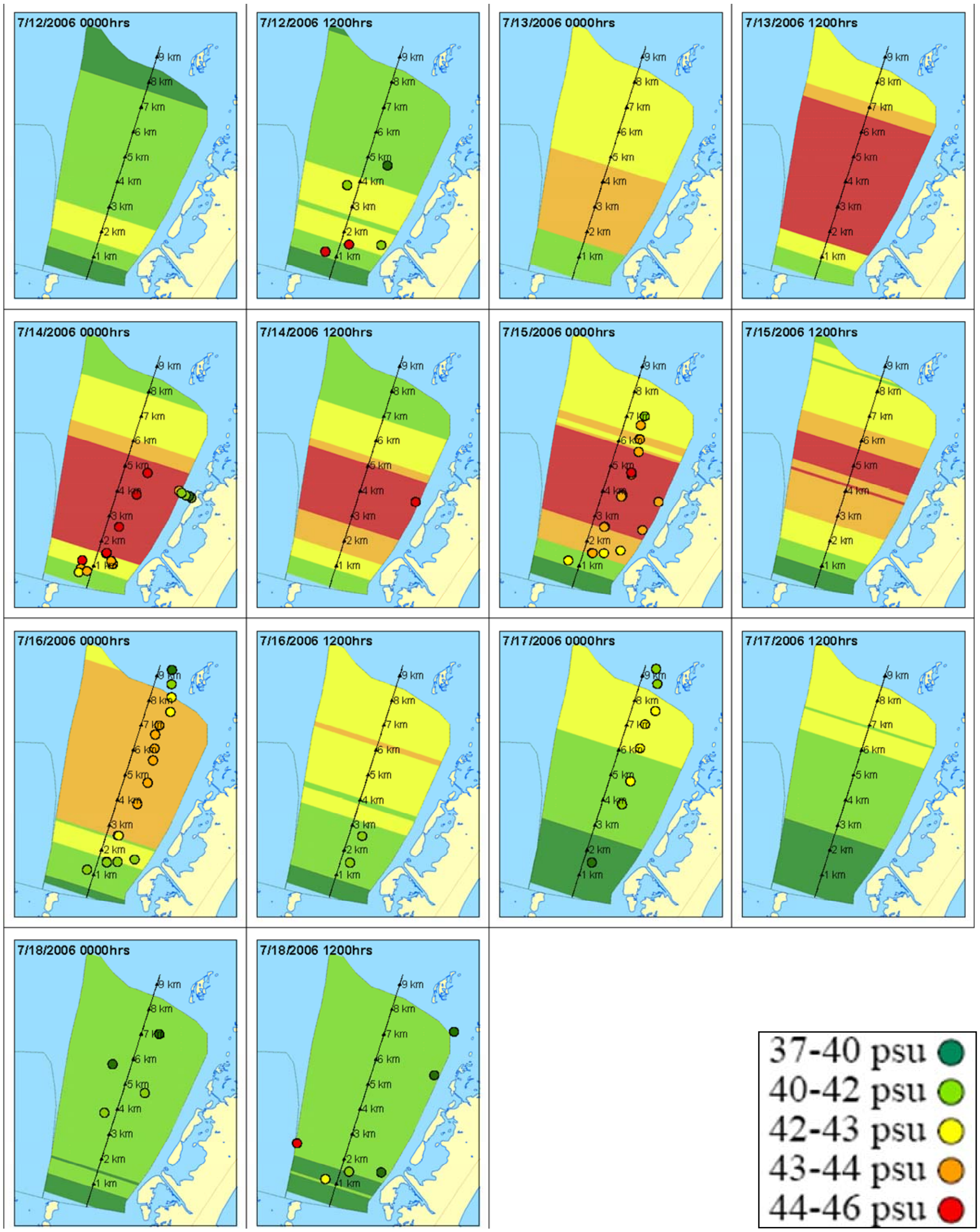


Figure 3.14. Bottom salinity in East SECCB study area from July 12 to 18, 2006.

.

CROSS VALIDATION OF RESULTS

Cross-validation is a common method that is used to evaluate the appropriateness of the adopted kriging model. This was performed using the "fictitious point" method (Delhomme, 1978), which involved removing one data point at a time from the data set and then using the remaining *n-1* points the estimate the removed point. The estimated and actual values of the data point were then compared with each other. A plot of the estimated salinities in the Bay versus the actual measured salinities is shown in Figure 3.15.

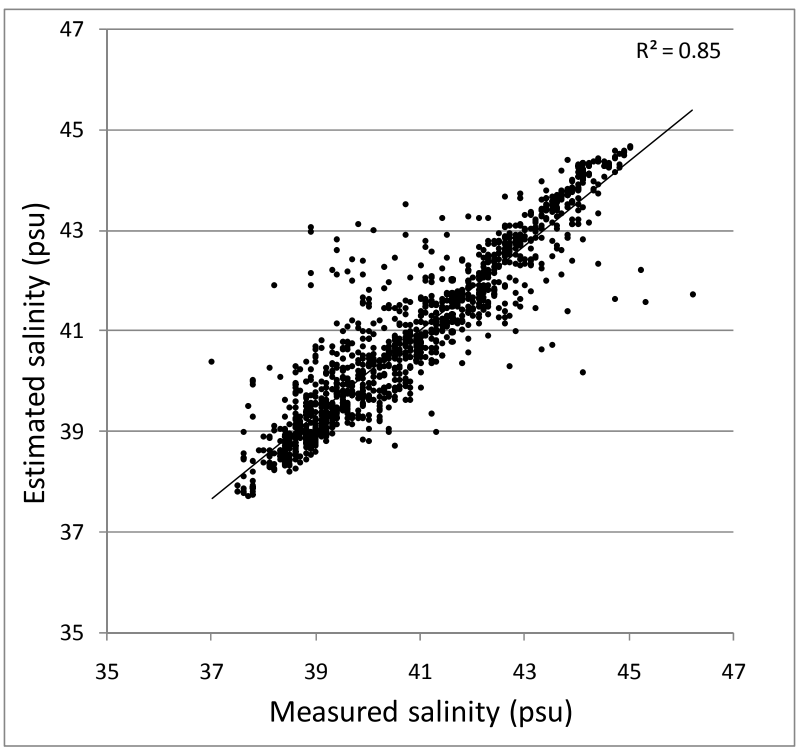


Figure 3.15. Cross validation results from space-time kriging.

A trend line is fitted and the r^2 value obtained is 0.85.

3.7 Conclusions for Chapter 3

The following observations are made from the visual interpretation of the kriging results:

1. The "life span" of a gravity current, in other words, the time from its descent into the Bay to its dissipation, was estimated to be on the order of 5 days (approximately one week).
2. The time it took for the gravity current to travel from the mouth of Laguna Madre (0 m) to the northern end of the study area (9000 m) was between 4 to 5 days. In other words the current speed was on the order of magnitude of 1 km/d.

From the cross-validation results, it was observed that the variogram models used in the kriging was able to explain 85% of the variability in the data. This performance was considered quite remarkable given the spottiness of the data and the fact that a statistical method was used to model natural data. The high performance could also be a result of the strength of the correlation in the salinity data over space and time. Unlike mineral deposits – which kriging is frequently applied to – salinity in water bodies are much less heterogeneous and changes over space and time are more gradual.

This study demonstrated that space-time kriging can facilitate the construction of a continuous space-time volume of salinity from fragments of data collected by multiple sensor networks. De Cesare et al's modified GSLIB proved to be a very useful tool for performing space-time kriging. By dissecting the space-time volume in regular time intervals, snapshots of the gravity current as it underwent stages of emergence, movement and dissipation were visualized. By interpreting the snapshots, the persistence of a gravity current in the Bay and its speed were estimated. The applicability of space-time kriging to data harvested by cyberinfrastructure has been demonstrated for Corpus Christi Bay. The potential of space-time kriging towards broader application lies in the fact that it is a statistical method and therefore general enough to be applied to a broad range of environmental variables and systems. This is important because the need for data synthesis is increasing as more and more environmental data are published via cyberinfrastructure. Space-time kriging allows environmental researchers to take advantage of the synergy provided by the unification of multiple data sources to obtain insights to environmental problems.

CHAPTER 4: EFFECTS OF WIND ON SALINITY STRATIFICATION IN SOUTHEAST CORPUS CHRISTI BAY

By Sin Chit To, Ben Hodges, Paul Montagna, Paula Kulis and David Maidment

4.1 Abstract

Hypoxia is the depletion of dissolved oxygen to levels harmful to aquatic organisms. It is a common environmental problem that affects many coastal water bodies in the United States. A major cause of hypoxia is density stratification in the water column – which imposes an energy barrier against the transfer of oxygen from the top to the bottom of the column. External forces like strong winds eliminate stratification and, in turn hypoxia, by vertically mixing the water column. However plume-tracking studies in southeast Corpus Christi Bay led by Ben Hodges from the University of Texas at Austin found stratification in a large part of the Bay immediately after a period of strong winds (> 4 m/s). Most of these winds blew from the southeast direction. This led to the development of hypothesis that certain winds actually induce stratification by forcing hypersaline waters from adjacent shallow bays (such as the Laguna Madre) into southeast Corpus Christi Bay. This paper presents a series of statistical tests conducted on data

recorded from 2005 to 2008 to 1) test the dominance of the gravity current theory in the Bay, and 2) identify what ranges of wind speed and direction are responsible for forcing hypersaline waters into the Bay. Results showed that wind patterns for different levels of stratification are significantly different from each other. Also the occurrence of strong winds (> 4m/s) blowing from the southeast direction is found to correlate positively with the level of stratification.

4.2 Introduction

SALINITY STRATIFICATION AND HYPOXIA

Hypoxia in aquatic systems refers to the situation when the dissolved oxygen concentration is below 2 mg/L (Dauer et al., 1992). Most organisms avoid, or become physiologically stressed, in waters with oxygen below this concentration (Diaz and Rosenberg, 1995). Hypoxia is caused by the exertion of oxygen demand by organic material under conditions of restricted water exchange and insufficient oxygen supply, often closely connected with density stratification of the water column (Pearson and Rosenberg, 1978 and Gray et al., 2002). Stratification reduces vertical turbulent mixing of heat, momentum, mass and constituents (Ralston and Stacey, 2005; Armenio and Sarker, 2002), and therefore limit the replenishment of dissolved oxygen in the bottom layer.

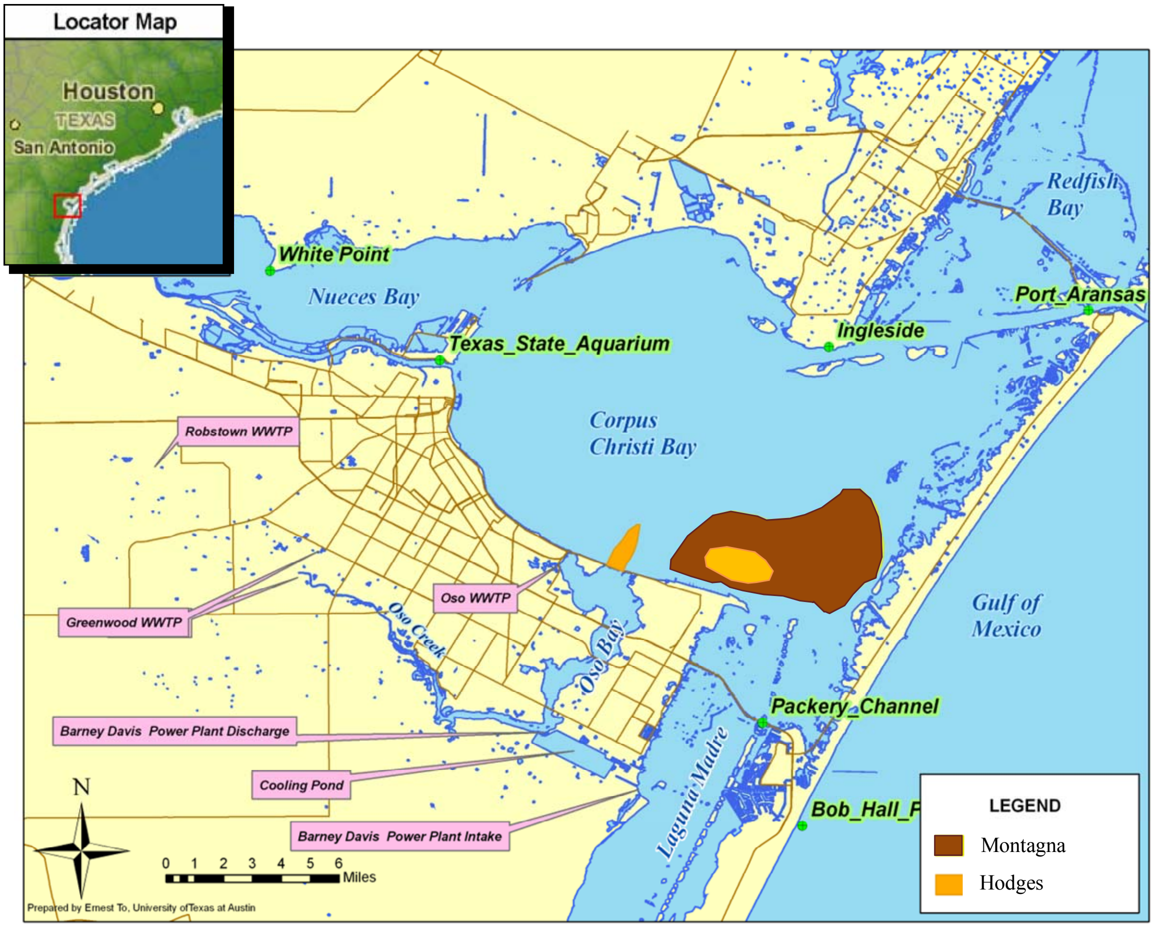
While hypoxia can occur naturally, it is often exacerbated by human impact such as excess nutrient enrichment of water bodies from point and non-point sources. More than half of the U.S. estuaries now experience natural or human-induced hypoxic conditions at some time each year and the frequency and duration of hypoxic events have increased over the last few decades (NOAA, 2007). According to Diaz and Rosenberg (2008), hypoxia problems are increasing world-wide, so understanding the mechanisms behind them and providing solutions is vital to the protection of the environment. Of the various mechanisms, investigating the cause of stratification takes priority because of its close relationship with hypoxia.

SALINITY STRATIFICATION IN CORPUS CHRISTI BAY

This research focuses on the hypoxia problem in Corpus Christi Bay, which is located along the Gulf of Mexico in south Texas (see Figure 4.1). Hypoxia and stratification in Corpus Christi Bay was first documented in 1988 (Montagna and Kalke, 1992) and later observed every summer (Martin and Montagna, 1995; Applebaum *et al.* 2005). Hypoxia occurs mainly in summer when temperature and evaporation are high and precipitation is low (Ritter and Montagna, 1999). Hypoxia causes a ten-fold reduction in benthic standing stock and diversity in the Bay.

Figure 4.1 shows the regions where stratification-induced hypoxia was observed in Corpus Christi Bay. These hypoxic regions were delineated by:

1. Dr. Paul Montagna, a marine biologist from the Harte Research Institute of Texas A&M University, Corpus Christi.
2. Dr. Ben Hodges, an environmental engineer at University of Texas at Austin.



Ernest Sin Chit To, CRWR

Figure 4.1. Hypoxia regions in Corpus Christi Bay delineated by Montagna and Hodges.

DESCRIPTION OF CORPUS CHRISTI BAY

The Corpus Christi Bay system is an urban estuary with complex hydrodynamic and water quality conditions. The bay is approximately circular in shape and has a diameter of approximately 13 miles. The average depth of the Bay is 9 feet, and is separated from the Gulf of Mexico by a barrier island, so that water circulation in the Bay is driven more by wind than by tides. Communication with the gulf is available via two channels:

1. Aransas Pass – a deep ship channel which connects Corpus Christi Bay to the Gulf of Mexico.
2. Packery Channel – a shallow channel that connects Laguna Madre to the Gulf.

Adjacent to Corpus Christi Bay is Laguna Madre which is a lagoonal system that is characterized by:

1. Shallow depth. They are about 3 to 4 feet deep (compared to 9 ft in Corpus Christi Bay).
2. Limited freshwater inflow. Most inflows are from surrounding surface runoff during storm events
3. Dense aquatic vegetation (Montagna, 1993).

Due to the shallowness of Laguna Madre and Corpus Christi Bay, as well as their limited interaction with the gulf, elevated salinity levels are found during periods in the summer when evaporation is high and precipitation is low. The typical salinity in the Gulf of Mexico ranges from 10 to 30 psu (practical salinity units). During the summer, salinity in Corpus Christi Bay can reach as high as 50 psu, while salinity in upper Laguna Madre can reach as high as 70 psu.

An interesting characteristic of hypoxia in Corpus Christi Bay is that a linkage between eutrophication and hypoxia has not been established (Applebaum et al, 2005) therefore hypoxia is predominantly correlated with salinity-induced stratification of the Bay. This study focuses on the causes of stratification in Corpus Christi Bay to provide insights to future attempts in modeling hypoxia in the Bay.

MECHANISMS FOR STRATIFICATION IN SOUTHEAST CORPUS CHRISTI BAY

Two mechanisms exist for the generation of stratification in water system

- *In situ* generation of stratification due to evaporation
- *Ex situ* generation of stratification due to introduction of gravity currents from Laguna Madre.

In situ generation of stratification: Evaporation

Heat from the sun evaporates surface water and increases its salinity. The denser saline water and sinks towards the bottom of the Bay – producing fingering patterns known as "salt-fingering". Over time, the dense water accumulates to form a bottom layer, resulting in stratification. The boundary between the bottom layer and the top layer is known as the pycnocline. Bottom oxygen levels are gradually depleted by benthic demand. At the same time, winds blowing over the Bay introduce mixing energy into the water column, which is then transferred down the water column. Mixing gradually erodes away the thickness of the bottom layer causing the gravity current to dissipate. Hypoxia happens when strong winds fail to dissipate the bottom layer before oxygen falls below hypoxic levels. The process is illustrated in Figure 4.2.

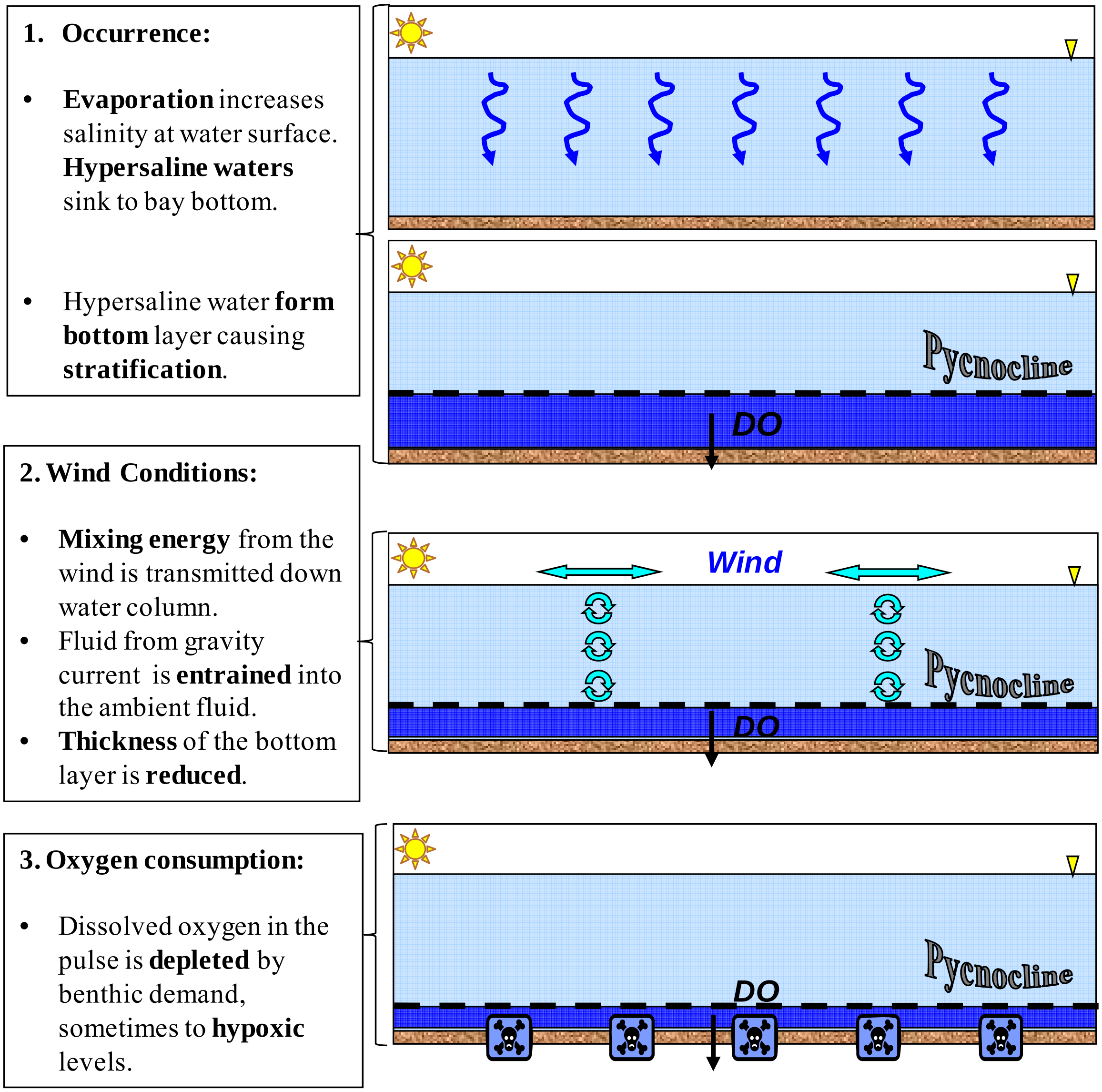


Figure 4.2. Effect of evaporation on stratification and hypoxia.

Southeast Corpus Christi Bay is susceptible to *in situ* generation of stratification because of two factors. First of all, bays in south Texas experience high evaporation rates and low precipitation during the summer (Applebaum and Montagna, 2004). The net evaporation rate during the summer months (that is the evaporation rate minus the precipitation rate) is 11 cm month^{-1} (Ward and Armstrong, 1997). Therefore conditions in south Texas are conducive for stratification. Secondly, southeast Corpus Christi Bay often experiences sustained periods of low water circulation (Powell *et al.*, 1997) and low wind speeds (Morehead *et al.*, 2002) in the summer. In addition, anthropogenic structures in the region (i.e., ship channel, Intracoastal Waterway, and a causeway across the Bay) have altered circulation patterns and created sluggish water circulation. As a result they impede the mixing and dissipation of the stratified layer. Applebaum and Montagna (2004) also observed that salinity stratification is significantly correlated with distance from the ship channel.

***Ex situ* generation of stratification: Gravity currents**

Ex situ generation of stratification is caused by hypersaline flows (also known as gravity currents) that enter southeast Corpus Christi Bay. This theory, also known as the gravity current theory, was put forward by Hodges and Kulis (Hodges, *et al.* 2008). The movement of these gravity currents is dominated by gravity, and flow of hypersaline waters occurs towards deeper areas by traveling along the bottom surface of the Bay. Dissolved oxygen transfer into the gravity current is limited and oxygen levels drop to hypoxic levels if the wind does not dissipate the gravity current first. However, wind also plays a creation role in stratification: dense aquatic vegetation in Laguna Madre impedes the movement of water and therefore a certain force is required to give a gravity current the momentum to overcome this resistance. Winds blowing in the right direction and speed provide this force. The process of *ex situ* generation of stratification is illustrated in Figure 4.3.

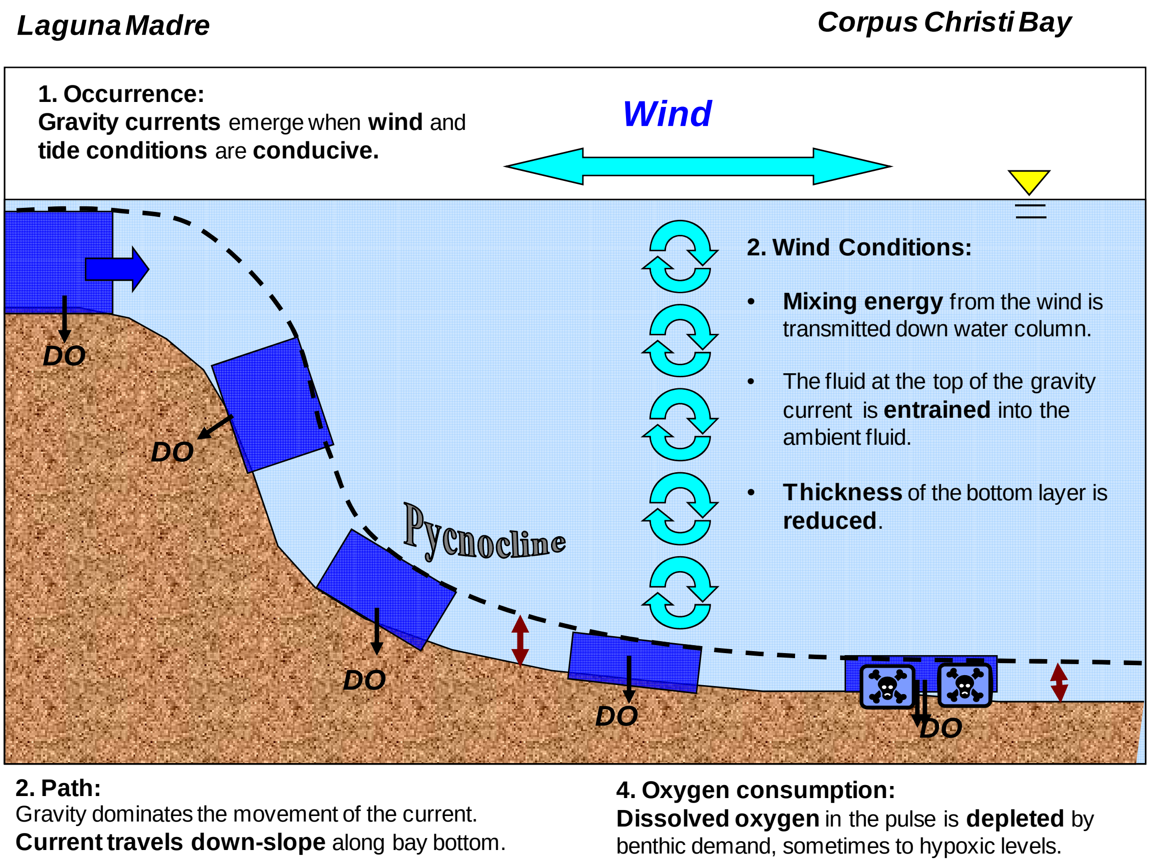


Figure 4.3. Gravity currents and hypoxia.

Summer evaporation increases the salinity in Laguna Madre and Oso Bay to levels above that of Corpus Christ Bay. Hypersaline waters can enter the Bay and cause stratification. Evidence of hypersaline flows have been found both downstream of Oso Bay (Hodges, et al, 2008) and Laguna Madre (Brower, *et al.* 2007). Therefore southeast Corpus Christi Bay is also susceptible to *ex situ* generation of stratification.

Since both *in situ* and *ex situ* mechanisms are possible, the question is therefore which mechanism dominates. Answering this question would not only help the understanding of the hypoxia problem but also the prediction of its occurrence, size and persistence.

DATA SOURCES IN CORPUS CHRISTI BAY

Several governmental agencies and academic institutions maintain environmental sensor networks in Corpus Christi Bay. The three networks that are relevant to this study are the HRI (Harte Research Institute) network, Hodges surveys (funded by the Texas Water Development Board), TCOON (Texas Coastal Ocean Observation Network). The locations of these sensors are shown in Figure 4.4. Table 4.1 provides a summary of the data they collect. A detailed discussion of each network is provided in the rest of this section.

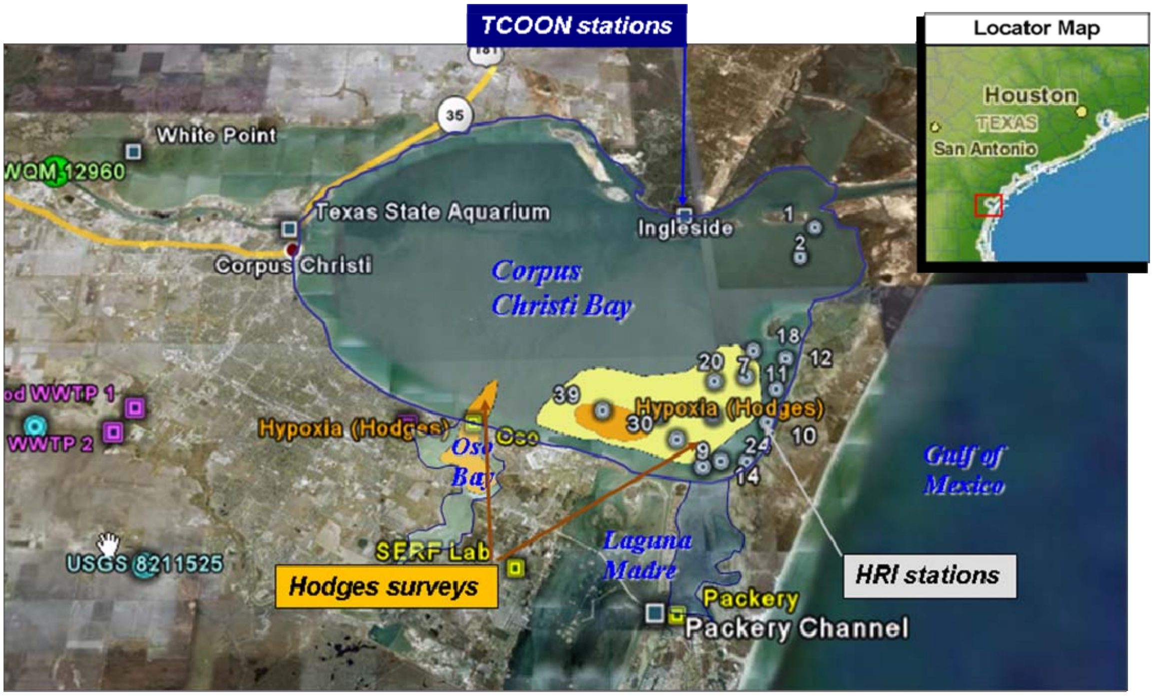


Figure 4.4. Sensor networks in Corpus Christi Bay

Table 4.1. Sensor networks in Corpus Christi Bay.

	Name of network	**Description**	**Data**
1.	HRI	Harte Research Institute	Collects water quality data
2.	Hodges surveys	Plume tracking studies performed by Dr. Ben Hodges of the University of Texas at Austin	Collects water data at high temporal frequencies to track the movement of hypersaline waters from Oso Bay and Laguna Madre.
3.	TCOON	Texas Coastal Ocean Observation Network	Collects wind and tide data

Harte Research Institute (HRI) network

In east Corpus Christi Bay, Dr. Paul Montagna and his team have collected water quality data over a wide area (~ 50 square kilometers) since the 1980s. The stations monitored are shown in grey in Figure 4.4. These data have undergone rigorous statistical analyses to characterize long-term spatial and temporal variability in hypoxia-related variables such as dissolved oxygen, salinity and temperature (Russell and Montagna, 2007). Relationships among the variables themselves have also been studied (Applebaum and Montagna, 2004, Russell and Montagna, 2007). Most data were collected at weekly to bi-weekly intervals during the summer months. Despite the extent of the data in time and space, it was noticed that the formation of stratification and hypoxic events typically occurred over a finer temporal scale (on the order of one week) than the sampling frequency. Therefore although the data captured the occurrence and extent of stratification, it was not sufficient to capture the process by which stratification developed.

Hodges plume tracking studies

In the region downstream of Oso Bay, Dr. Ben Hodges and his team have performed short-term intensive data collection over small areas (~ 1 to 2 square miles) to capture the formation and dissipation of stratification (Hodges and Furnans, 2006, Hodges, et al., 2008) (see Figure 4.4). They found that hypersaline flows, also known as gravity currents, from Oso Bay are associated with stratification and hypoxia in the area.

In 2006, Dr. Hodges performed an intensive plume tracking study in the Bay just downstream of Laguna Madre (Brower, *et al.* 2007). Salinity, oxygen and temperature data were collected every 12 hours. The study captured a gravity current as it emerged from the mouth of Laguna Madre and descended into the Bay.

Texas Coastal Ocean Observing Network

The Texas Coastal Ocean Observation Network (TCOON) is a set of observing stations located along the Texas coast that measure water level, wind speed, direction and a variety of environmental conditions, including water and air temperature. These stations collect data continuously at six-minute intervals. Six TCOON stations are located around Corpus Christi Bay, as shown in Figure 4.4. Historical data going back to 1993 are available for most parameters at these stations.

Conclusions from data sources

The juxtaposition of the various sensor networks in the Bay indicates that southeast Corpus Christi Bay has the highest concentration of data. More importantly, it contains data from both intensive short-term study from Hodges and long-term extensive monitoring from Montagna. The synthesis of these two data sets offers a unique opportunity to find both the cause of stratification and its occurrence. This research focuses on southeast Corpus Christi and insights acquired from this area can potentially be applied to other parts of the Bay.

4.3 Role of wind in stratification and hypoxia

DIFFERENCES IN EFFECT OF WIND BETWEEN *IN SITU* AND *EX SITU*

In situ and *ex situ* stratification differ not only in the source of hypersalinity, but also in the effect of the wind.

- In *in situ* generation of stratification, the only effect of the wind is the dissipation of the bottom layer. Therefore wind reduces stratification and does not cause it.
- In *ex situ* generation of stratification, the wind plays two roles. On one hand, it can dissipate the bottom layer. On the other hand, if it blows in a certain direction and at certain speed, it can provide an external force that pushes hypersaline waters into Corpus Christi Bay. Therefore even though in general wind reduces stratification, certain types of wind can cause it.

To determine whether the *in situ* or *ex situ* mechanism dominates, it is necessary to find out whether trends exist between wind patterns and stratification events.

TRENDS OBSERVED IN HISTORICAL DATA

Between 1/1/2005 and 9/13/2007, Paul Montagna and Ben Hodges collected salinity profiles in the southeast Corpus Christi Bay for a total of 37 days. Salinity profiles from each sampling day were mapped. Figures 4.5, 4.7, and 4.9 are three examples of these maps. For each sampling day, the percentage of profiles that were stratified (i.e. those having density differences between the top and bottom layer > 2 g/L) was calculated. The wind pattern surrounding each sampling day was plotted as wind vectors. Each wind vector represents the speed and direction of one wind observation. A wind vector can be visualized as an arrow where the tail arises from its source and the length of the shaft represents its speed – the longer the shaft, the faster the wind. The data for the wind patterns were measured by TCOON (Texas Coastal Ocean Observation Network. The wind patterns corresponding to Figures 4.5, 4.7, and 4.9 are shown in Figures 4.6, 4.8 and 4.10. Comparison of the wind history and stratification fraction yielded insights as to what classes of winds cause stratification. The following is a discussion of the trends shown in Figures 4.5 to 4.10.

August 16, 2005 – Stratification in majority of experimental area

Figure 4.5 shows the stratification state and wind for August 16, 2005. This is a day where 100% of the profiles were stratified. Figure 4.6 showed that in the 7 days prior to this date, a consistent strong wind (~ 6 m/s) was blowing from the southeast direction.

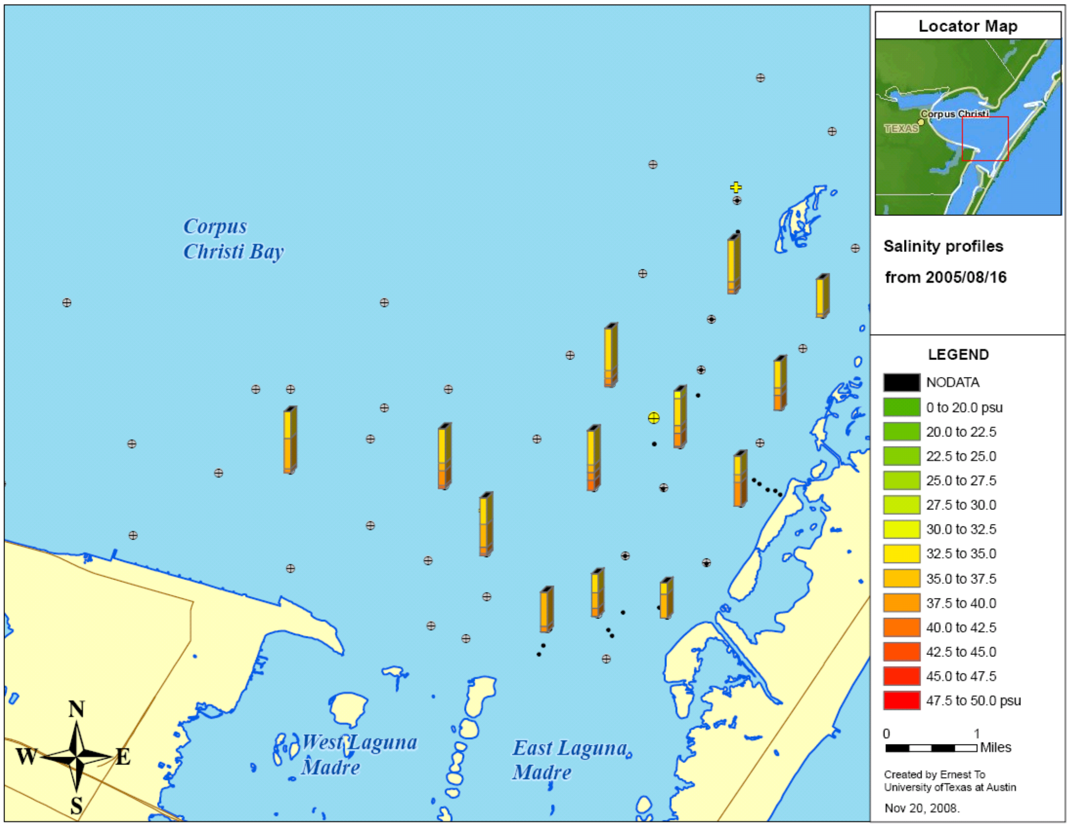


Figure 4.5. Salinity profile on Aug 16, 2005 showing stratification in most of Southeast Corpus Christi Bay.

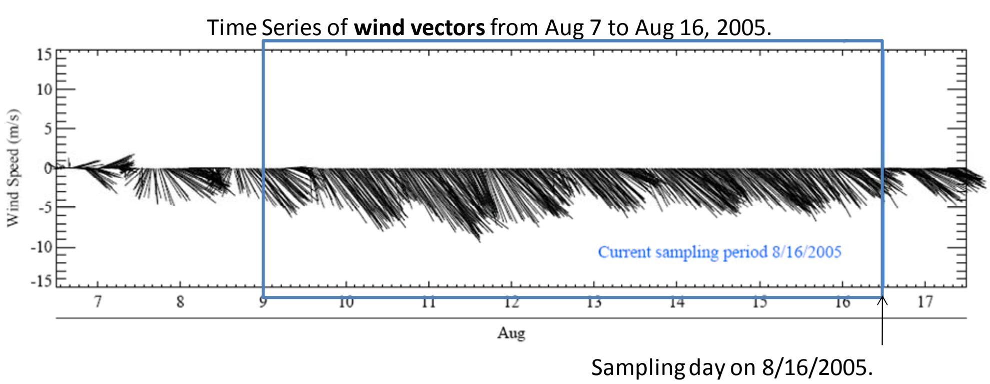


Figure 4.6. 10-day wind history surrounding Aug 16, 2005 (from Aug 7, 2005 PM to Aug 17, 2008 AM).

June 28, 2005 – No stratification in experimental area.

Figure 4.7 shows the stratification state and wind for June 28, 2005. This is a day where 0% of the profiles were stratified. Figure 4.8 showed that in the 8 days prior to this date, the wind was blowing consistently from the east.

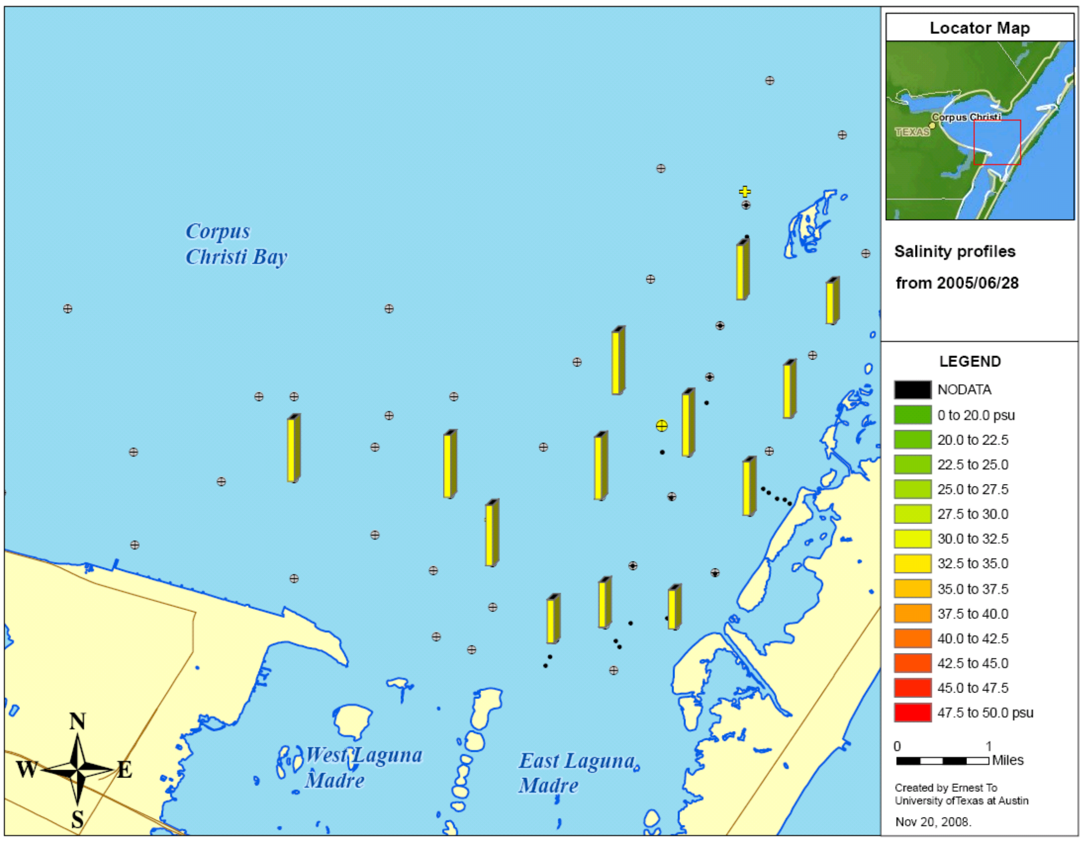


Figure 4.7. Salinity profile on June 28, 2005 showing no stratification in Southeast Corpus Christi Bay.

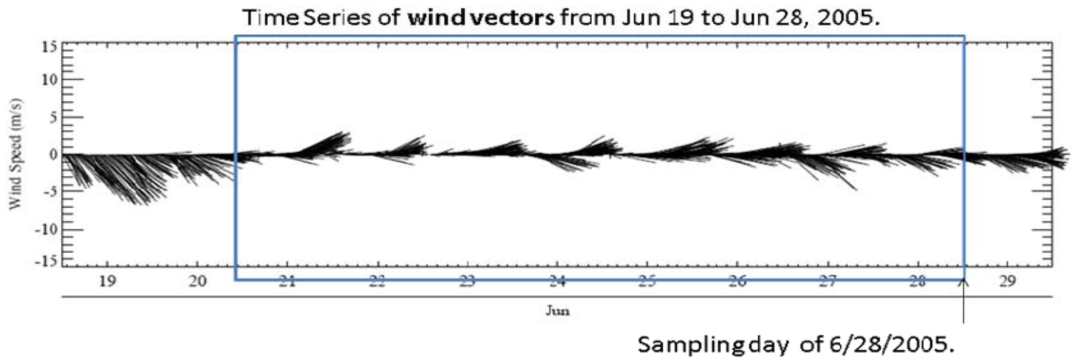


Figure 4.8. 10-day wind history surrounding June 28, 2005 (from June 19, 2005 PM to June 29, 2008 AM).

July 27, 2005 – Stratification in part of the experimental area.

Figure 4.9 shows the stratification state and wind for July 27, 2005. This is a day where 62% of the profiles were stratified. Figure 4.10 showed that in the 9 days prior to this date, wind first blew from the northeast for 2 days, then from an assortment of directions for 2 days, then finally from the southeast for 5 days.

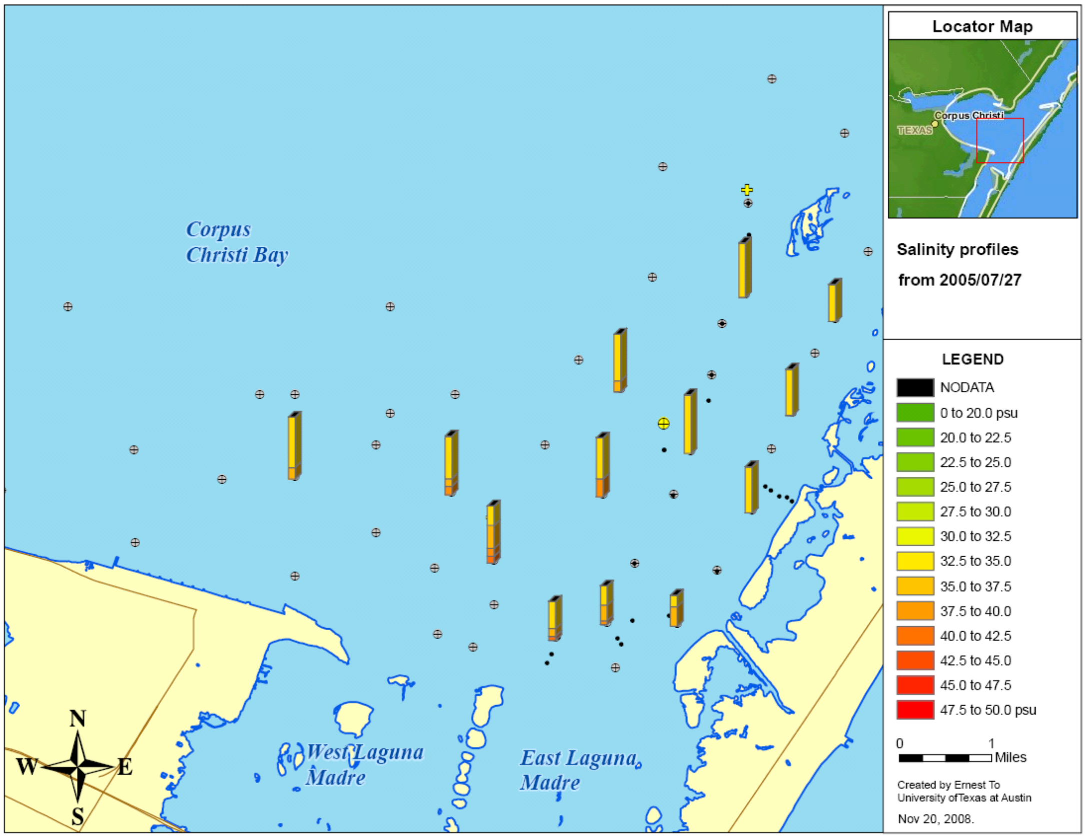


Figure 4.9. Salinity profile on July 27, 2005 showing stratification in part of southeast Corpus Christi Bay

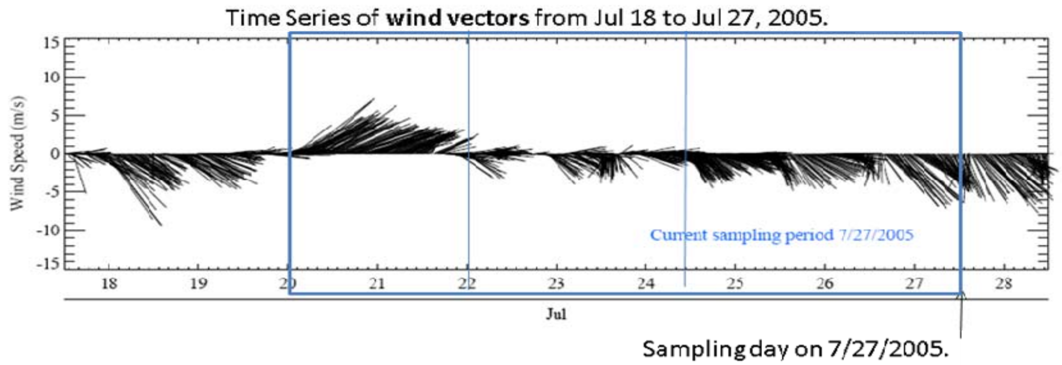


Figure 4.10. 10-day wind history surrounding July 27, 2005 (from July 18, 2005 PM to July 28, 2008 AM)

In all three cases, the stratification fraction in the Bay is observed to increase with the increase in occurrence of winds from the southeast. To relate this observation to the underlying mechanism in the Bay, a valve model hypothesis was posed.

VALVE MODEL HYPOTHESIS

Under *ex situ* generation of stratification, winds of certain speed and direction provide the external force to push gravity currents into the Bay. The resulting mechanism is like operating a valve on a hose. Some winds turn the valve on and release a gravity current, other winds leave the valve off and release no gravity currents (see Figure 4.11). A wind event is defined as a wind that turns the valve on and cause stratification. More wind events leads to stratification in a larger portion of the Bay.

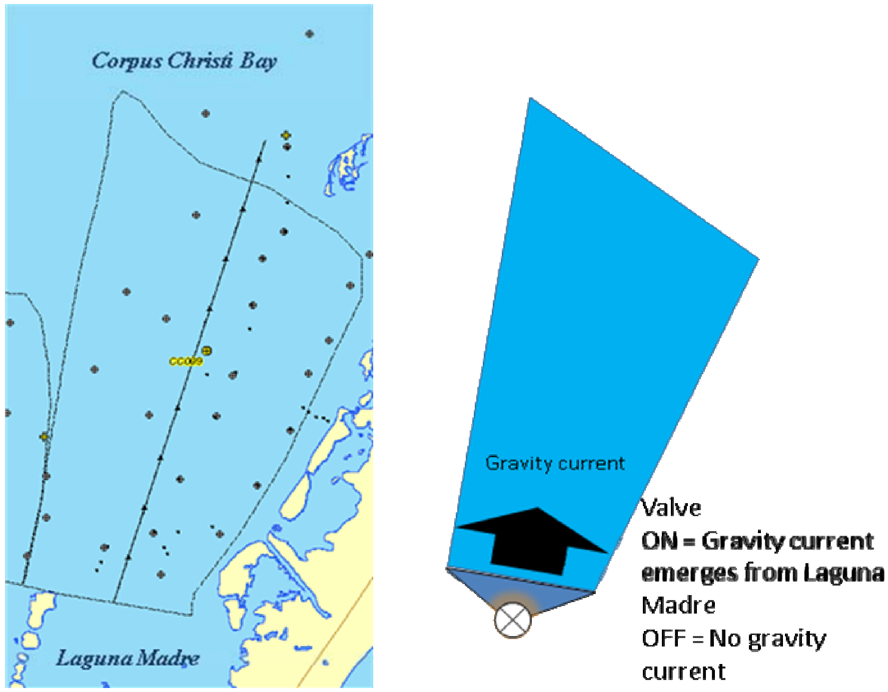


Figure 4.11. Concept of a valve model.

4.4 Selection of experimental area

In order to study the underlying mechanism behind stratification in southeast Corpus Christi bay, it was necessary to define an experimental area for hypotheses testing and modeling. Scientific insights gathered from analyzing data from this area can then be transferred to other parts of the Bay. Two requirements were set for selecting the experimental area:

1. The area has to be small enough to comprise only one water system. Analyzing an area that spans multiple water systems introduces noise into the analysis. From an *ex situ* stratification stand point, this means the area can only be influenced by one source of hypersaline waters. For *in-situ* stratification, area defined must be

small enough so that precipitation and evaporation are uniform over the entire area.

2. The area has to be big enough to be well-populated by sensors to provide sufficient data for analysis.

USING BATHYMETRY TO DEFINE ZONES OF INFLUENCE

The selection of a suitable experimental area is met by considering the gravity current theory. Because the movement of gravity currents is dominated by gravity, they flow down the slope in a manner similar to runoff in a watershed. Bathymetric features such as depressions and ridges limit the advancement of gravity currents. Ridges act as barriers to block the movement of gravity currents. Depressions act as sinks that trap the gravity current. In 2007, NOAA published new bathymetry data in Corpus Christi Bay as part of its SIFT (Short-Term Inundation Forecasting for Tsunami) project (NOAA, 2007b). These bathymetry data indicate the presence of the following significant terrain features in south Corpus Christi Bay (see Figure 4.12):

A. An underwater ridge that runs from north to south into the chain of islands that divides the region downstream of Laguna Madre into east and west portions.

B. An underwater ridge that runs from northeast to southwest through the center of the Bay.

C. A depression that lies directly north of the western portion of Laguna Madre.

D. A depression that lies to the northwest of Laguna Madre.

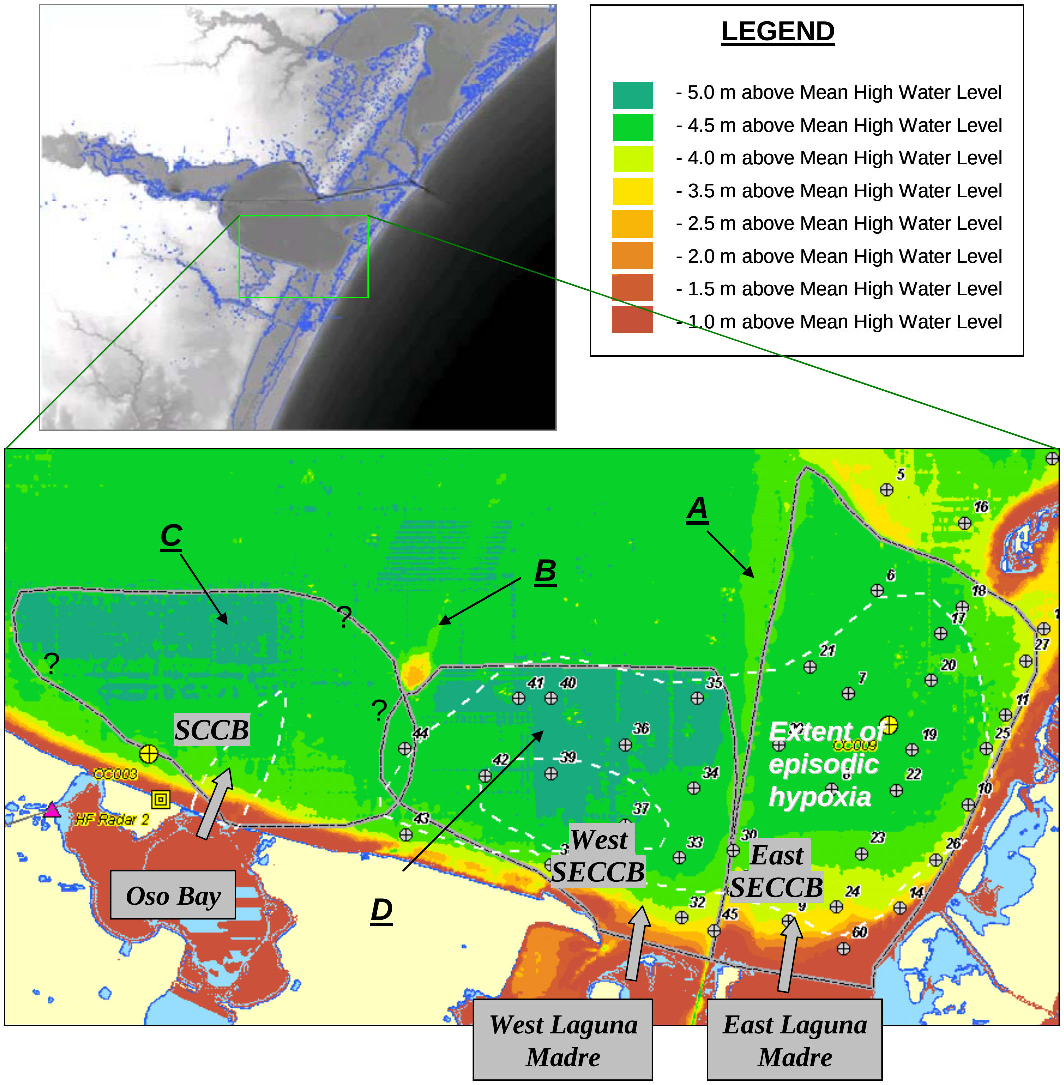


Figure 4.12. Terrain features in bay bottom compartmentalize effects of gravity currents.

The underwater terrain features compartmentalize the Bay into underwater basins, each influenced by a different shallow bay. These basins are:

1. East Southeast Corpus Christi Bay Area (East SECCB), which is influenced by East Laguna Madre;
2. West Southeast Corpus Christi Bay Area (West SECCB), which is influenced by West Laguna Madre;

3. South Corpus Christi Bay Area (SCCB), which is influenced by Oso Bay.

One of the most interesting observations from the bathymetry data is that the union of the East and West SECCB areas matches with the extent of the episodic hypoxic area observed by Paul Montagna using his HRI stations (see white dash line in Figure 4.12). This is an extra piece of evidence that supports the *ex situ* mechanism for stratification. Because East SECCB contains the most sensors, and therefore the most data-rich, it is selected as the experimental area to test the dominance of *in situ* vs. *ex situ* generation of stratification.

4.5 Wind-stratification analysis

Plume tracking data collected Dr. Hodges in 2006 demonstrated that *ex situ* stratification can occur in southeast Corpus Christi Bay. However there still remained the question of whether it is the dominant cause of stratification in the area. Because none of the sensor networks in the Bay collected salinity data at high enough temporal resolution, it is not possible to examine the processes leading to each recorded stratification event. In the absence of high-resolution temporal salinity data, one looks into the driving force behind *ex situ* stratification, i.e. the wind, to test if it indeed is the dominant cause. Therefore, high-resolution wind data from TCOON is the key data element to analyze the cause of stratification. This section describes how a series of hypothesis tests 1) demonstrate the relationship between wind and stratification and 2) identify certain categories of winds that contribute to stratification.

DATA PREPARATION

The analysis of the relationship between wind and stratification requires 1) a metric for quantifying the degree of stratification in the Bay and 2) a means of summarizing the wind patterns in the Bay. This section presents the methodology for doing the above.

Quantification of stratification events

The stratification intensity for a salinity profile collected in the experimental area of southeast Corpus Christi Bay can be quantified using sigma-t (σ_t). Sigma-t is the difference in density (measured in parts per thousand, or g/L) between the top and bottom of the water column.

$$\sigma_t = \text{top density (g/L) - bottom density (g/L)} \qquad \text{(Equation 4.1)}$$

The conversion of salinity (in practical salinity units) to density (g/L) can be performed using the Equation 4.2:

$$\rho = 1000 + 0.761 \times \text{salinity} \qquad \text{(Equation 4.2)}$$

Equation 4.2 is a linear approximation of the UNESCO's regression equations for salinity to density conversion (UNESCO, 1980). The EPA criterion for strong stratification is σ_t > 2 g/L (EPA, 2008). Using this criterion, a given salinity profile is classified as either a stratified profile or a non-stratified profile depending on whether its σ_t is greater or less than 2 g/L. To summarize the state of stratification for a given sampling day, the stratification fraction is calculated by dividing the number of profiles that are stratified with the total number of profiles (Equation 4.3).

$$\text{stratification fraction} = \frac{\text{number of stratified profiles}}{\text{total number of profiles collected}} \qquad \text{(Equation 4.3)}$$

The stratification fraction can also be expressed as a percentage value. For instance, when 15 out of 20 profiles collected in the Bay are stratified (i.e. stratification fraction = 0.75), the Bay is experiencing 75% stratification.

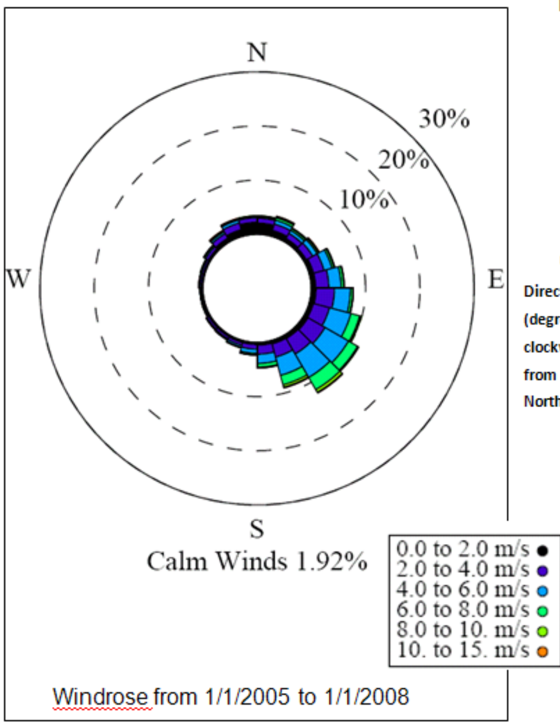
Summarizing wind histories as wind roses

To summarize wind patterns in the Bay, the prior 5-day wind histories for each sampling days is plotted as a wind rose. A 5-day wind history is used because the time of travel of a gravity current in the study area is about 5 days according to the TWDB plume tracking study by Hodges.

Categorizing wind vectors into wind classes

The wind vectors are categorized into 144 wind classes, which represents 24 directional intervals and 6 speed intervals (see Figure 4.13). The directional intervals are 15 degree intervals starting from 0 to 360 degrees (measured clockwise from the North). The speed intervals are 0 to 2 m/s, 2 to 4 m/s, 4 to 6 m/s, 6 to 8 m/s, 8 to 10 m/s and 10 to 15 m/s. All observed wind speeds between 1/1/2005 to 1/1/2008 were between 0 to 15 m/s.

Each wind rose is a bivariate (i.e. speed and direction) probability distribution of the wind. Figure 4.13 shows the average wind rose of winds measured from 1/1/2005 to 1/1/2008. On the right side of the figure shows how this wind rose is represented in tabular format. Each cell in the table is the probability of each wind class. The probability of each wind class is calculated as the sum of the durations of each wind that falls within the wind class, and then divided by the total period of 5 days. This bivariate distribution of the wind was used in subsequent statistical tests to validate the valve model hypothesis.



f(θ,V) — V Wind Speed (m/s); θ Direction (degrees clockwise from North)

	0 to 2	2 to 4	4 to 6	6 to 8	8 to 10	10 to 15
0 to 15	2.06%	0.92%	0.17%	0.13%	0.07%	0.05%
15 to 30	1.07%	1.24%	0.73%	0.46%	0.18%	0.11%
30 to 45	0.65%	1.09%	0.64%	0.36%	0.10%	0.03%
45 to 60	0.50%	1.36%	0.93%	0.29%	0.10%	0.06%
60 to 75	0.78%	2.00%	1.52%	0.38%	0.11%	0.02%
75 to 90	0.80%	2.43%	2.27%	0.50%	0.04%	0.01%
90 to 105	0.88%	3.39%	2.95%	0.49%	0.03%	0.01%
105 to 120	0.74%	3.08%	4.09%	1.67%	0.26%	0.01%
120 to 135	0.83%	3.61%	5.01%	1.94%	0.19%	0.01%
135 to 150	0.70%	3.30%	5.00%	2.60%	0.59%	0.11%
150 to 165	0.63%	2.51%	3.51%	1.80%	0.47%	0.11%
165 to 180	0.49%	1.67%	1.71%	0.69%	0.11%	0.03%
180 to 195	0.37%	0.87%	0.75%	0.28%	0.04%	0.01%
195 to 210	0.36%	0.52%	0.37%	0.09%	0.00%	0.00%
210 to 225	0.32%	0.44%	0.16%	0.02%	0.00%	0.00%
225 to 240	0.31%	0.41%	0.11%	0.01%	0.00%	0.00%
240 to 255	0.26%	0.21%	0.05%	0.00%	0.00%	0.00%
255 to 270	0.21%	0.21%	0.07%	0.01%	0.00%	0.00%
270 to 285	0.26%	0.31%	0.11%	0.04%	0.00%	0.00%
285 to 300	0.25%	0.33%	0.14%	0.07%	0.02%	0.01%
300 to 315	0.41%	0.54%	0.26%	0.12%	0.04%	0.01%
315 to 330	0.71%	0.99%	0.39%	0.15%	0.08%	0.09%
330 to 345	1.23%	1.29%	0.56%	0.15%	0.08%	0.09%
345 to 360	2.20%	0.85%	0.15%	0.11%	0.06%	0.02%

Figure 4.13. Wind roses as probability distributions.

Grouping of wind roses by stratification fraction

The wind roses for the sampling days are ranked by the stratification fraction. Figure 4.14 shows an excerpt of the ranked series.

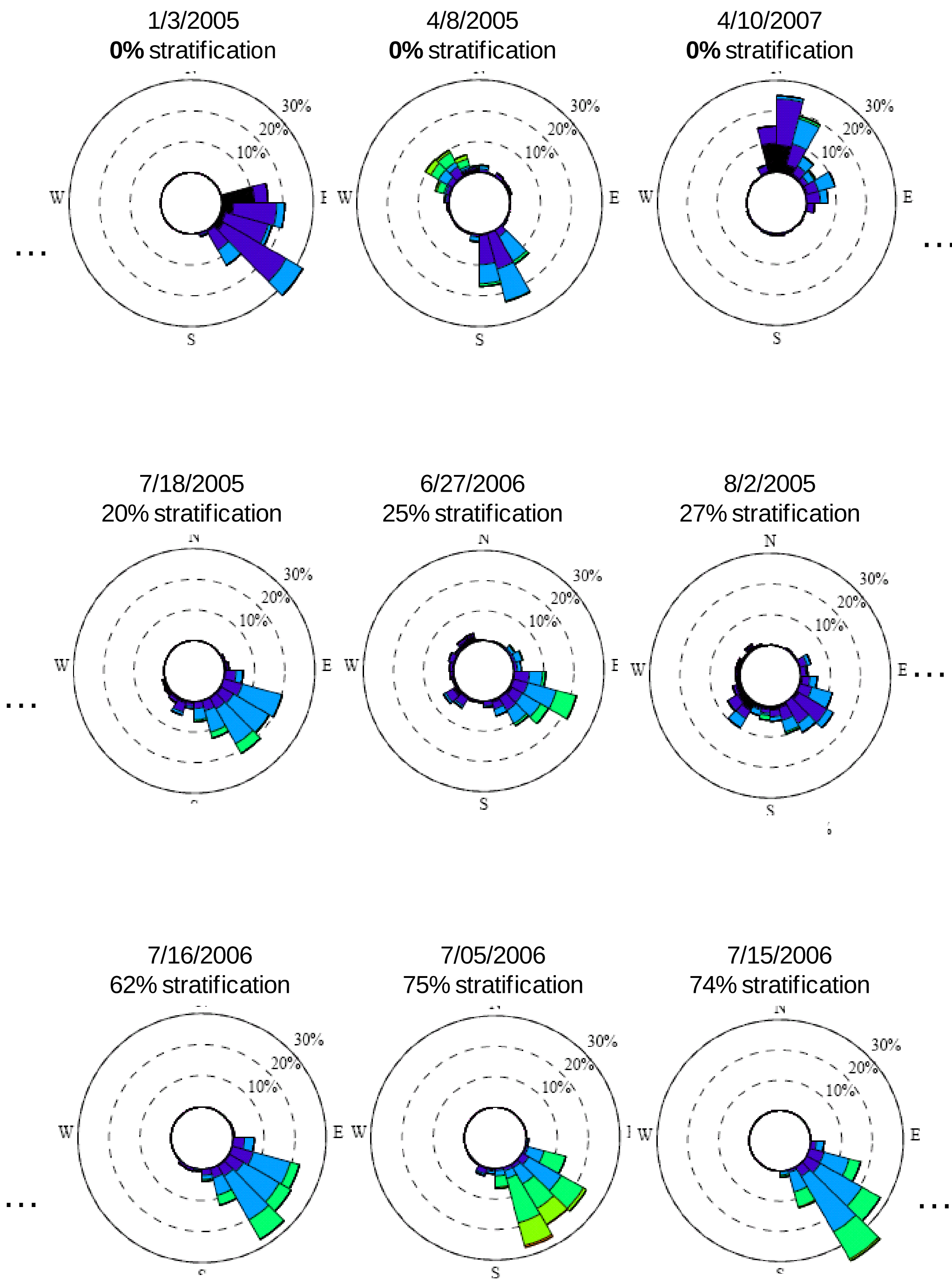


Figure 4.14. Wind roses of sampling days ranked by stratification fraction.

To eliminate the apparent noise in the wind distributions the wind roses are grouped into three categories: 0% stratification, >0% to 50% stratification and >50% to 100% stratification. The averaged wind roses of each of the three categories are shown in Figure 4.15. Figure 4.15 also shows the average wind rose for the entire period of record (1/1/2005 to 1/1/2008).

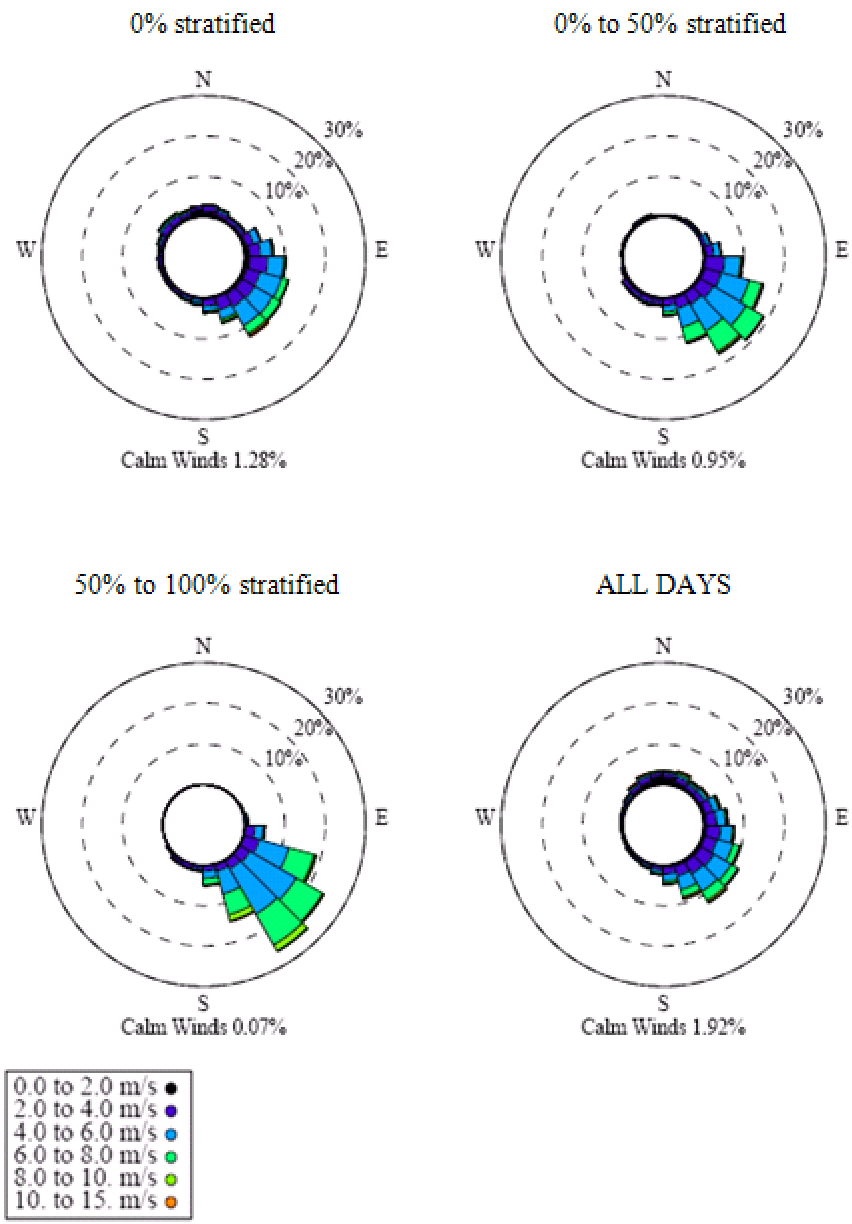


Figure 4.15. Average wind roses for sampling days that have 0% stratification, >0% to 50% stratification and >50% to 100% stratification and for all days from the entire period of record from 1/1/2005 to 1/1/2008.

From Figure 4.15, one can observe that the percentage of southeasterly winds with speeds greater than 4 m/s increases with the stratification fraction. However, this is not sufficient to conclude that strong, southeasterly winds cause stratification. Recall that stratification is usually found during the summer months (June, July, August) when precipitation is low and evaporation is high. If stratification is the result of seasonality and not of the wind (in other words *in situ* stratification dominates over *ex situ* stratification) then the wind distributions shown in Figure 4.15 is merely a reflection of the wind distribution converging towards the typical summer wind distribution. It is therefore necessary to address the following three questions:

1. Are wind distributions from different levels of stratification significantly different from each other?
2. Which classes of winds exhibit significant positive correlation with the stratification fraction?
3. Is the variability in the stratification fraction better explained by the season or by the wind distribution?

To address these questions, a series of hypothesis tests is performed.

HYPOTHESIS TESTING

Do wind patterns change significantly as stratification increases? Are wind distributions from different levels of stratification significantly different from each other?

Chi-square tests are applied to verify that the differences between pairs of wind distributions from different stratification fractions are statistically significant. The chi-square test is a comparison of entire sample distributions and therefore more powerful than just the comparison of means (as represented by broad brush tests such as ANOVA and MANOVA). The chi-square test can conclusively determine that the differences in wind distributions are significant.

The chi-square test is a goodness of fit test that measures of how far a sample distribution deviates from a theoretical distribution (Zar, 1999). It is applied by Wallis and Griffith (1997) to compare wind roses predicted by models with those derived from wind measurements. For this analysis, the chi-square test is used to establish that the wind distributions for 0% stratification, >0% to 50% stratification and >50% to 100% stratification are all significantly different from one other.

The chi-square statistic is calculated using Equation 4.4

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i} \qquad \text{(Equation 4.4)}$$

where

χ^2 is the chi-square statistic;

f_i is the frequency, or number of counts, observed in the wind class i;

$\hat{f}_i$ is the frequency expected in class i if the null hypothesis is true; and,

k is the number of wind classes.

Since there are 144 wind classes (24 wind direction ranges * 6 wind speed ranges) the number of degrees of freedom for chi-square test is 144-1 = 143. Equation 4.4 shows that larger differences between two distributions result in a higher chi-square statistic.

The chi-square test is used to compare the following pairs of wind roses:

1. 0% stratification vs. >0% to 50% stratification
2. >0% to 50% stratification vs. >50% to 100% stratification
3. 0% stratification vs. >50% to 100% stratification

For each comparison, the following null and alternate hypotheses are stated.

- Null hypothesis: H_0: Wind distribution #1 is *no different* from Wind distribution #2
- Alternate hypothesis: H_A: Wind distribution #1 is *significantly different* from Wind distribution #2.

The significance level of $p < 0.05$ is used to reject the null hypothesis. However, because the test involves three pairwise tests, the significance level was adjusted using the Bonferroni correction to 0.05/3 = 0.017. The Bonferroni correction states that if an experimenter is testing *n* dependent or independent hypotheses on a set of data, then one way of maintaining the familywise error rate is to test each individual hypothesis at a statistical significance level of 1/*n* times what it would be if only one hypothesis were tested.

In all the above three cases (0% stratified vs. >0 to 50% stratified; >0 to 50% stratified vs. >50 to 100% stratified; and, 0% stratified vs. >50% to 100% stratified) the null hypothesis is rejected at a significance level of 0.017. Therefore wind distributions from days with higher stratification fraction are significantly different from days with lower stratification fraction.

Which classes of winds exhibit significant positive correlation with the stratification fraction?

Having demonstrated that wind patterns from different levels of stratification are significantly different from each other, the next challenge is to identify which wind classes change with increasing stratification fraction. A wind event is a wind class that increases with stratification fraction. At first glance, this can be investigated by applying a logistic regression of the stratification fraction versus the 144 wind classes (see Equation 4.5). From the regression, wind classes that exhibit significantly positive slopes are identified as wind events.

$$\mathrm{logit(p)} = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_{144} X_{144}$$

Equation 4.5

where p is the stratification fraction;

logit() is the logistic transformation function;

X_j is the occurrence of a wind class, j, within a 5-day period – quantified as a percentage of the entire period;

b_0 is the intercept of the regression; and,

b_1 to b_{1144} are the slopes of the 144 wind classes.

Logistic regression is applied instead of linear regression because the stratification fraction is a probability of obtaining stratified salinity profiles on a given sampling day. Therefore it has a value between 0 and 1. Logistic regression transforms the probability into a parameter called the ***logit*** which ranges from $-\infty$ to $+\infty$ (Equation 4.6). This prevents the regression from generating erroneous models that predict outside the 0 to 1 range (Helsel and Hirsch, 1995).

$$\text{logit(p)} = \log(\frac{p}{1-p}) \qquad \text{Equation 4.6}$$

where p is the stratification fraction.

However, because there exists only a total of 37 observations of stratification fraction from the 37 sampling days, regressing 144 wind classes with Equation 4.5 results in overspecification. For this reason, it is necessary to group the wind classes into broader categories to reduce the number of explanatory variables.

The strategy is therefore to first perform one-on-one regression between the occurrence of each wind class, j, and stratification fraction using the form:

$$\mathrm{logit(p)} = b_{0j} + b_{1j} X_j$$

Equation 4.7

Where p is the stratification fraction;

X is the occurrence of a wind class, j, within a 5-day period – quantified as a percentage of the entire period;

b_{0j} is the intercept of the wind class, j; and,

b_{1j} is the slope of the wind class, j.

Slope (b_{0j}) from the logistic regression is tested to see if it is positive and significantly different from zero. Based on the one-on-one regression, the wind classes are grouped into two groups:

1. A wind-event group that comprises of wind classes that possess significant, positive relationships with stratification fraction (X_{ON}).
2. A non-wind event group that comprises of all other classes (X_{OFF}). .

A final regression is performed to 1) confirm that there exists a significant positive relationship between the wind event group and stratification fraction; and, 2) evaluate the ability of wind events to explain the uncertainty in stratification fraction.

The confirmation regression equation has the form:

$$\text{logit(p)} = b_0 + b_1 X_{ON} + b_2 X_{OFF}$$

Equation 4.8

Where p is the stratification fraction as a fraction;

X_{ON} is the occurrence of a wind event occurring within a 5 day period;

X_{OFF} is the occurrence of a non-wind event occurring within a 5 day period;

b_0 is the intercept,

b_1 is the slope of the wind event.

b_2 is the slope of the non-wind event.

Because X_{OFF} is the complement of X_{ON}, (i.e. $X_{OFF} = 1\text{-}X_{ON}$) the regression simplifies to Equation 4.9.

$$\text{logit(p)} = b_0 + b_1 X_{ON}$$

Equation 4.9

Logistic regression on wind classes

When performing individual regressions with Equation 4.7, the maximum likelihood estimation (MLE) method is used instead of ordinary least squares (OLS) because the distribution of the stratification fraction and its logit are not normal. This results in the violation of the OLS assumption of homoskedasticity. To implement MLE, a search algorithm is programmed to choose the slope and intercept parameters that maximize the log-likelihood of the resulting regression model.

The log-likelihood is a measure of the likelihood that the observed data would be produced from a given set of slopes and hence is a measure of the performance of the regression model. The greater the log-likelihood,value, the better the performance. The equation of the log-likelihood is given in Equation 4.10.

$$l_j = \sum_{i=1}^{n} (y_{ij} \cdot \ln[\hat{p}_i] + (N_i - y_{ij}) \ln[1 - \hat{p}_i]) \qquad \text{Equation 4.10}$$

where

l_j is the log-likelihood for the regression of wind class, j;

y_i is the number of stratified salinity profiles on a particular sampling day, i;

N_i is the total number of salinity profiles on a particular sampling day, i;

n is the total number of sampling days, which is 37;

$\hat{p}_i$ is the expected probability of finding a stratified salinity profile on a particular sampling day, i.

Evaluation of logistic regression results

Overall likelihood ratio tests are performed to identify slopes that are statistically significant (Helsel and Hirsch, 1995). The overall likelihood ratio compares whether the logistic regression model is better than an intercept only model (i.e., one that has a slope of zero – which indicates no correlation) by taking the difference between the log-likelihoods of the regression model and the intercept-only model.

The following null and alternate hypotheses of the test are stated:

Null hypothesis: Stratification is independent of the occurrence of wind in the wind class.

$$H_0: \; b_j = 0$$

Alternate hypothesis: Stratification is dependent of the occurrence of wind in the wind class.

$$H_A: \; b_j \neq 0$$

where b_j is the slope coefficient for wind class, j, from the one-on-one logistic regression (recall Equation 4.8).

The log-likelihood ratio between the logistic regression model and the pure-intercept model is given by Equation 4.11:

$$l_{rj} = 2 \cdot (l_{cj} - l_{0j}) \qquad \text{Equation 4.11}$$

Where

l_{rj} is the overall log-likelihood ratio for wind class, j.

l_{cj} is the log-likelihood of the logistic regression model for wind class, j.

l_{0j} is the log-likelihood of the pure-intercept model for wind class, j.

The overall likelihood ratio is approximated by a chi-square distribution with 1 degree of freedom, which is the number of additional coefficients the regression model possesses over the pure-intercept model. The null hypothesis is rejected at a significance level of 0.05. However, because this hypothesis test is performed for 144 wind class, the significance level for each wind class was corrected to 0.05/144 = 0.00035.

The signs of the slope coefficients for the 144 wind classes and their significance are displayed in Table 4.2. Table cells that are highlighted in black are wind classes that are tested to have significantly positive slopes. Table cells that are highlighted in grey are wind classes that have positive slopes which are not significantly different from zero. The table cells that are in white are those that have negative or zero slopes. Cells that have "Not-a-Number" (NaN) values are wind classes with no occurrences.

Table 4.2. Slope coefficients calculated from logistic regression for the 144 wind classes.

		Velocity (meters per second)					
		0 to 2	2 to 4	4 to 6	6 to 8	8 to 10	10 to 15
Degrees clockwise from North	0 to 15	-	-	-	-	NaN	NaN
	15 to 30	-	-	-	-	-	NaN
	30 to 45	-	-	+	+	NaN	NaN
	45 to 60	-	-	-	-	-	NaN
	60 to 75	-	-	-	-	-	NaN
	75 to 90	-	-	-	-	NaN	NaN
	90 to 105	-	-	-	-	-	NaN
	105 to 120	-	+	-	+	-	NaN
	120 to 135	-	-	+	+	+	-
	135 to 150	-	-	+	+	+	-
	150 to 165	-	-	+	+	+	+
	165 to 180	-	-	-	+	+	NaN
	180 to 195	-	-	-	-	-	NaN
	195 to 210	-	-	-	-	NaN	NaN
	210 to 225	-	-	-	NaN	-	-
	225 to 240	-	-	-	NaN	NaN	NaN
	240 to 255	-	-	-	NaN	NaN	NaN
	255 to 270	-	-	-	NaN	NaN	NaN
	270 to 285	-	-	-	-	NaN	NaN
	285 to 300	-	-	-	-	-	NaN
	300 to 315	-	-	-	-	-	NaN
	315 to 330	-	-	-	-	-	NaN
	330 to 345	-	-	-	-	-	NaN
	345 to 360	-	-	-	-	NaN	NaN

LEGEND

NaN	No occurrence
-	Negative
+	Not significantly positive
+	Significantly positive

The wind classes with significantly positive slopes are considered wind events that turn the valve ON, and the rest of the wind classes turn the valve OFF. Figure 4.16 shows that the wind events constitute winds that blow from N120 ^{o}W to N165 ^{o}W at a velocity between 4 and 8 m/s.

		Velocity (meters per second)					
		0 to 2	2 to 4	4 to 6	6 to 8	8 to 10	10 to 15
Degrees clockwise from North	0 to 15	OFF	OFF	OFF	OFF	NaN	NaN
	15 to 30	OFF	OFF	OFF	OFF	OFF	NaN
	30 to 45	OFF	OFF	OFF	OFF	NaN	NaN
	45 to 60	OFF	OFF	OFF	OFF	OFF	NaN
	60 to 75	OFF	OFF	OFF	OFF	OFF	NaN
	75 to 90	OFF	OFF	OFF	OFF	NaN	NaN
	90 to 105	OFF	OFF	OFF	OFF	OFF	NaN
	105 to 120	OFF	OFF	OFF	OFF	OFF	NaN
	120 to 135	OFF	OFF	ON	ON	OFF	OFF
	135 to 150	OFF	OFF	ON	ON	OFF	OFF
	150 to 165	OFF	OFF	ON	ON	OFF	OFF
	165 to 180	OFF	OFF	OFF	OFF	OFF	NaN
	180 to 195	OFF	OFF	OFF	OFF	OFF	NaN
	195 to 210	OFF	OFF	OFF	OFF	NaN	NaN
	210 to 225	OFF	OFF	OFF	NaN	OFF	OFF
	225 to 240	OFF	OFF	OFF	NaN	NaN	NaN
	240 to 255	OFF	OFF	OFF	NaN	NaN	NaN
	255 to 270	OFF	OFF	OFF	NaN	NaN	NaN
	270 to 285	OFF	OFF	OFF	OFF	NaN	NaN
	285 to 300	OFF	OFF	OFF	OFF	OFF	NaN
	300 to 315	OFF	OFF	OFF	OFF	OFF	NaN
	315 to 330	OFF	OFF	OFF	OFF	OFF	NaN
	330 to 345	OFF	OFF	OFF	OFF	OFF	NaN
	345 to 360	OFF	OFF	OFF	OFF	NaN	NaN

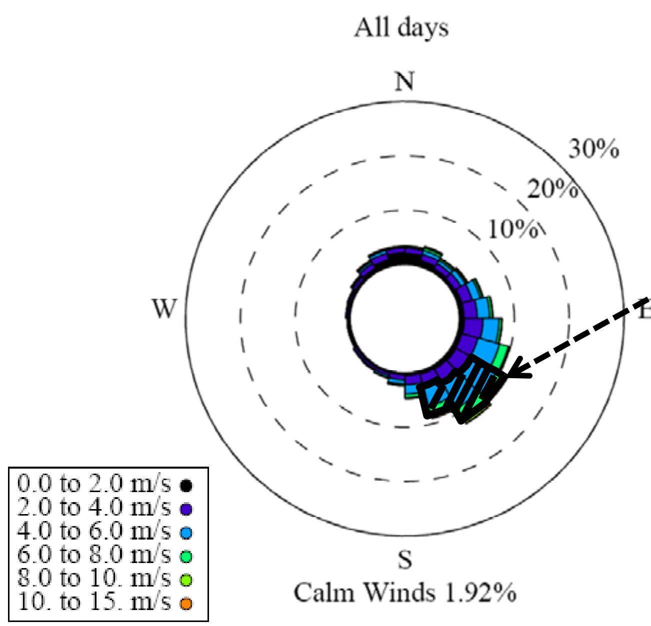


Figure 4.16. Wind classes identified as wind events that cause stratification.

Confirmation analysis

To confirm that 1) the wind-event group is positively correlated with stratification fraction and 2) quantify the amount of uncertainty in the stratification fraction that is explained by wind events, the following logistic regression is performed:

$$\text{logit(p)} = b_0 + b_1 X_{ON} \qquad \text{(recall Equation 4.9)}$$

where p is the stratification fraction;

X_{ON} is the probability of a wind event occurring within a 5 day period;

b_0 is the intercept,

b_1 is the slope of the wind event.

The resulting slope of the regression calculated by MLE is 4.5 and the intercept is -2.8. The graph of the regression model is shown in Figure 4.17. The log-likelihood of the regression model is calculated to be -181 using Equation 4.10 while the log-likelihood of a pure intercept model (i.e. one that does not depend on X_{ON}) is -210. The overall log-likelihood ratio, l_r, is calculated to be 58 (Equation 4.11). Using the chi-square test with 1 degree of (based on the number of additional coefficients between the regression model and the pure intercept model), one can reject the null hypothesis that the regression model in Equation 4.9 is no different from the pure intercept model at a significance level of 0.05. Therefore wind event group is significantly and positively correlated with stratification fraction.

A measure of the amount of uncertainty explained by the logistic regression model is the likelihood-R^2 or McFadden's R^2 (Hensel and Hirsch, 1995). The likelihood-R^2 is calculated as the proportion of log-likelihood (Equation 4.12).

$$R^2 = 1 - \frac{l}{l_0} \qquad \text{Equation 4.12}$$

Where

l is the log-likelihood of the logistic regression model.

l_0 is the log-likelihood of the pure-intercept model.

The likelihood-R^2 is 0.14 which means the wind events explain a minor portion of the uncertainty in stratification. This is not unexpected because stratification also depends on other factors such as the presence of hypersaline waters in Laguna Madre, the dissipation effect of winds as well as the spatial locations of the salinity profiles.

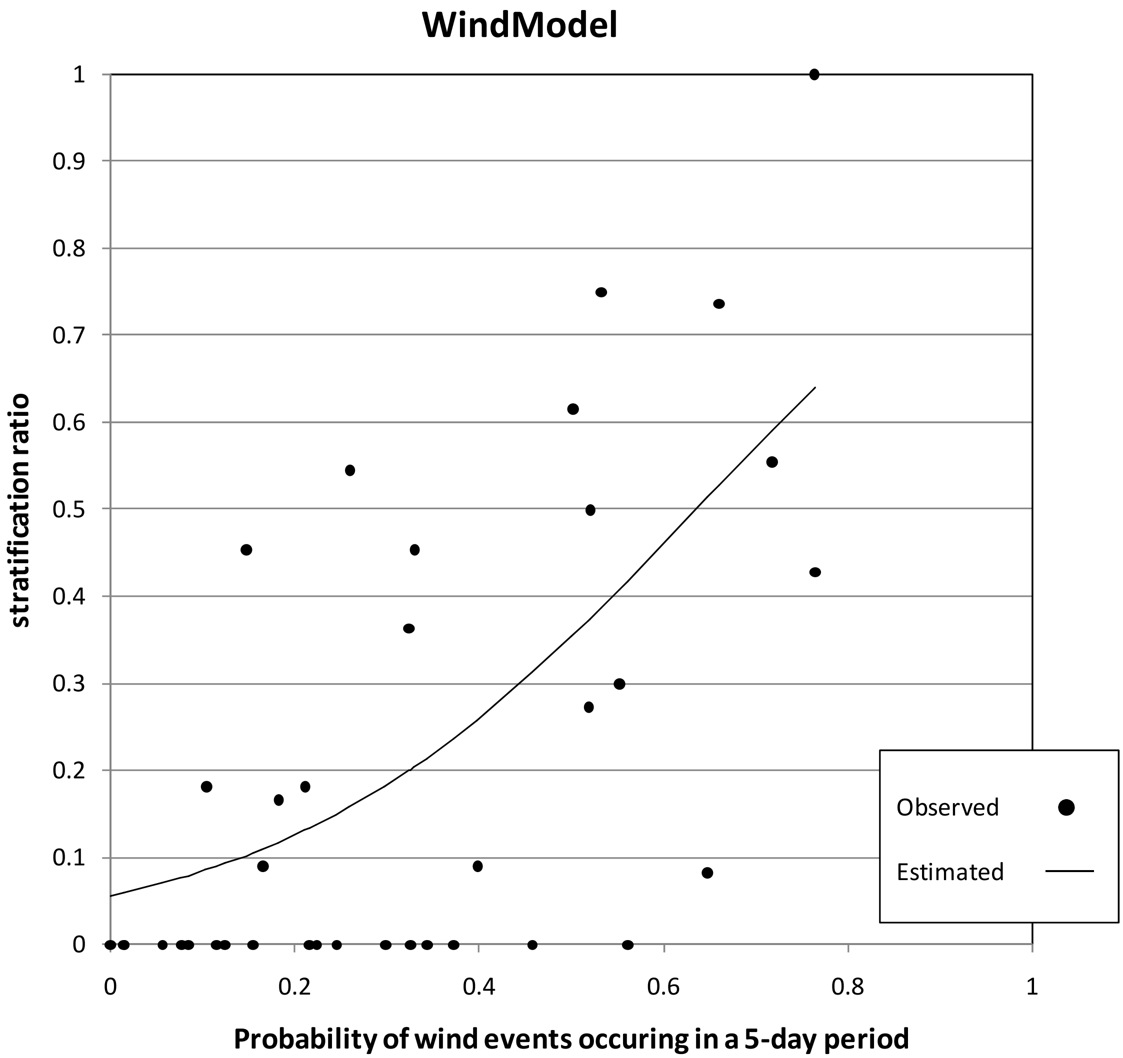


Figure 4.17. Graph comparing estimated stratification fraction (p_{est}) with measured stratification fraction

Is the variability in the stratification fraction better explained by the season or by the wind distribution

Logistic Regression models

To examine whether the different levels of stratification are a reflection of seasonal variations unrelated to wind or a result of wind forcings, a series of four logistic regressions were carried out to investigate the ability of seasonality to explain the stratification fraction. A binary variable, S, is defined where S equals 1 when the sampling day falls in the months, June, July and August. For all other months, S equals 0. The following four models are proposed:

Intercept model: $\mathrm{logit(p)} = b_0$

Season model: $\mathrm{logit(p)} = b_0 + b_1 S$

Wind model: $\mathrm{logit(p)} = b_0 + b_1 X_{ON}$

Wind-Season model: $\mathrm{logit(p)} = b_0 + b_1 S + b_2 X_{ON} + b_3 S X_{ON}$

The Intercept model acts as the benchmark for comparing the rest of the three models. It proposes that neither wind nor season has an effect on stratification. The Season model proposes that only the season has an effect on stratification. The Wind model proposes that only the wind has an effect on stratification. Lastly, the Wind-Season model proposes that both wind and season have an effect on stratification.

Using maximum likelihood estimation (MLE), parameters from the regression are obtained as follows:

Intercept model: $\text{logit(p)} = -0.975$

Season model: $\text{logit(p)} = -15.0 + 14.1S$

Wind model: $\text{logit(p)} = -2.83 + 4.45X_{ON}$

Wind-Season model: $\text{logit(p)} = -19.1 + 16.4S + 1.63X_{ON} + 2.66SX_{ON}$

Table 4.3 reports 1) the log-likelihoods, l, calculated using Equation 4.10; 2) the number of explanatory variables; and 3) Akaike Information Criteria (AIC) for each model. The log-likelihood is the natural log of the likelihood that the model will produce the estimated data. Better models tend to possess higher log-likelihoods. To penalize against models that artificially inflate their log-likelihoods by introducing redundant variables, the AIC is introduced. AIC a statistic that measures model error (given by the log-likelihood -l) and a penalty for too many variables, the number of explanatory variables, k (see Equation 4.13) (Hirsch and Helsel). Smaller AICs indicate better models.

$$AIC = -l + k \qquad \text{Equation 4.13}$$

Table 4.3 Log-likelihoods, number of explanatory variables and Akaike Information Criteria for the four logistic regression models.

	Overall log-likelihood (l)	Number of explanatory variables (k)	Akaike Information Criteria (AIC = -l + k)
Intercept Model	-210.1	1	211.1
Season Model	-205.5	2	207.5
Wind Model	-180.9	2	183.0
Wind-Season Model	-179.3	4	183.3

Conclusion from the Intercept, Season, Wind and Wind-Season models

From Table 4.3, the performance of the models is ranked as follows (from best to worst based on AIC): 1. Wind Model, 2. Wind Season Model, 3. Season Model, and 4. Intercept Model. Despite using an additional explanatory variable, the Wind-season model offers no significant improvement over the Wind model. The Wind Model also offers significantly better explanation than the Season Model. From the performance of the four models, one can conclude that the effect of seasonal variation on stratification is relatively insignificant compared to the effect of the wind. Therefore results from the analysis strongly support the claim that stratification is caused predominantly by the *ex situ* mechanism rather than *in situ* mechanism.

4.6 Conclusions for Chapter 4

This paper investigates the cause of stratification in Corpus Christi Bay by employing a series of statistical tests to determine whether stratification happened *in situ* (i.e. by evaporation) or *ex situ* (by gravity currents). The valve model hypothesis is proposed and tested. The valve model hypothesis states that under *ex situ* generation of stratification, more wind events leads to stratification in a larger portion of the Bay. The hypotheses tests conducted in this study demonstrated that wind patterns have a significant effect on stratification in the study area. The occurrence of winds that blow from the southeasterly direction at a speed between 4 and 8 m/s has a significant, positive correlation with stratification fraction. These winds have been identified as events that introduce gravity currents into the Bay according to the valve model concept. The effect of the wind on stratification is proven to be much more significant than seasonal effects. Therefore results from the analysis strongly support the claim that stratification is caused predominantly by the *ex situ* mechanism rather than *in situ* mechanism.

Despite showing significant, positive correlation between wind events and stratification, the occurrence of wind only explains a minor portion of the variation in stratification pattern. This is not unexpected because stratification also depends on other factors such as the presence of hypersaline waters in Laguna Madre, the dissipation effect of winds as well as the spatial locations of the salinity profiles. A full development of the valve hypoxia model that takes into account the other factors is needed in order to properly explain the stratification pattern in the Bay.

CHAPTER 5: MODELING THE EFFECTS OF WIND ON HYPOXIA IN SOUTHEAST CORPUS CHRISTI BAY

By Sin Chit To, Ben Hodges, Paul Montagna, Paula Kulis and David Maidment

5.1 Abstract

Hypoxia is a common environmental problem that affects many coastal water bodies in the United States. Hypoxia is related to density stratification in the water column – which imposes an energy barrier against the transfer of oxygen from the top to the bottom section of the column. In southeast Corpus Christi Bay, strong southeasterly winds force hypersaline waters from adjacent shallow bays (e.g. Laguna Madre) into the Bay, thereby causing density stratification. A simple plug flow model (the valve hypoxia model) was constructed to simulate the fate and transport of gravity currents. The model simulates the release of gravity currents during wind events is simulated by a valve mechanism. Comparison of model results with historical data from 2005 to 2008 showed that wind-driven gravity currents offer a compelling explanation of hypoxia patterns in southeast Corpus Christi Bay. A simple plug flow model (or the valve hypoxia model) was constructed to simulate the fate and transport of gravity currents that emerge during wind events. Comparison of model results with historical data from 2005 to 2008 showed that wind-driven gravity currents offer a compelling explanation of hypoxia patterns in southeast Corpus Christi Bay.

5.2 Introduction

HYPOXIA AS AN ENVIRONMENTAL PROBLEM

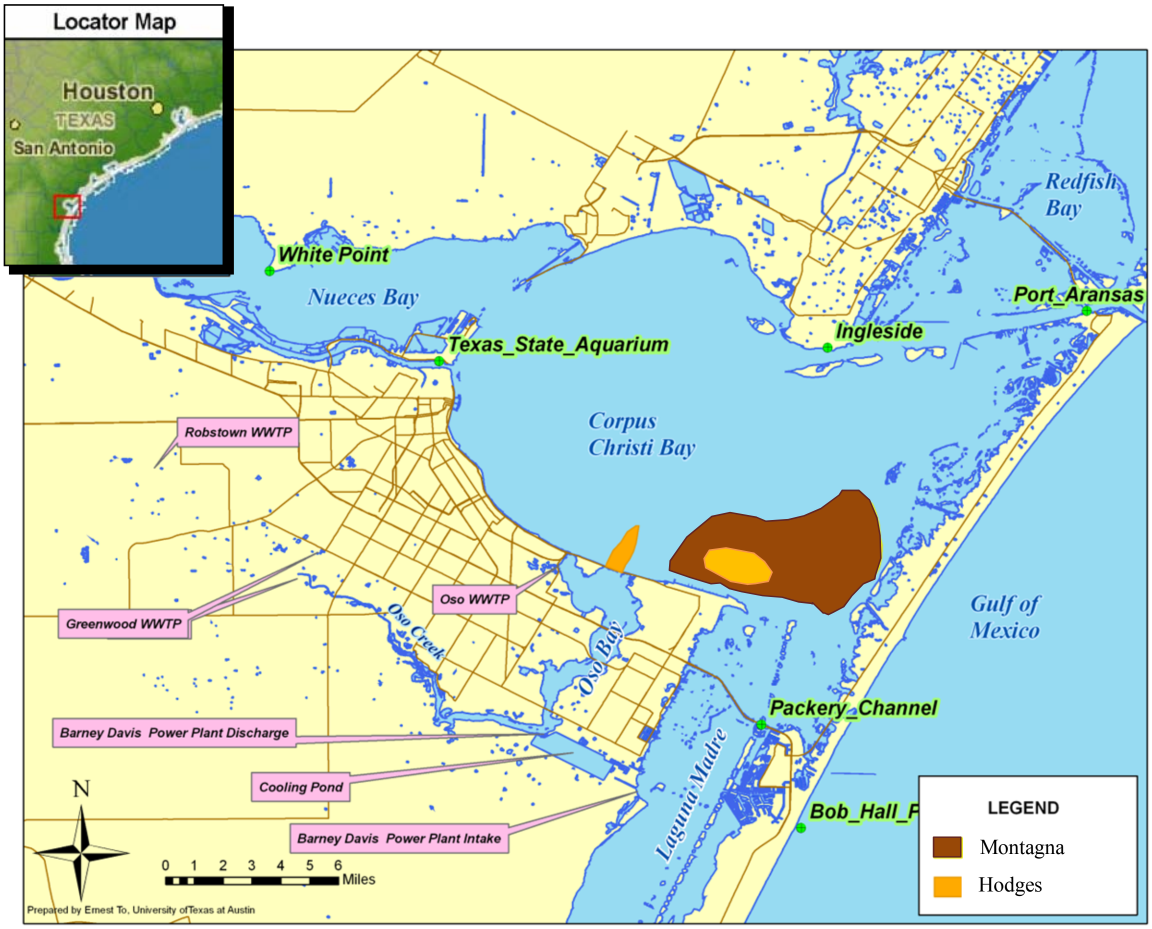
Hypoxia in aquatic systems refers to waters where the dissolved oxygen concentration is below 2 mg/L (Dauer et al., 1992). Most organisms avoid, or become physiologically stressed, in waters with oxygen below this concentration (Diaz and Rosenberg, 1995). Hypoxia affects commercial harvests and the health of ecosystems. While hypoxia can occur naturally, it is also a symptom of environments stressed by human impact such as from excess nutrient enrichment from point and non-point sources. Over half of the U.S. estuaries now experience natural or human-induced hypoxic conditions at some time each year and evidence suggests that the frequency and duration of hypoxic events have increased over the last few decades (NOAA, 2007). According to Diaz and Rosenberg (2008), hypoxia problems are increasing world-wide, so providing and understanding of the mechanisms and possible solutions is important.

HYPOXIA IN CORPUS CHRISTI BAY

Hypoxia in Corpus Christi Bay, Texas is first documented in 1988 (Montagna and Kalke, 1992) and later observed every summer (Martin and Montagna, 1995; Applebaum *et al.* 2005). Hypoxia in Corpus Christi Bay results in about a ten-fold reduction in benthic standing stock and diversity. Unlike other systems along the Gulf of Mexico, e.g. the Louisiana coast a linkage between eutrophication and hypoxia has not been established in

Corpus Christi Bay (Applebaum et al, 2005). Instead, hypoxia is found to be correlated with salinity-induced stratification of the Bay, which occurs in summer when temperature and evaporation are high and precipitation is low (Ritter and Montagna, 1999). Stratification causes hypoxia by reducing vertical turbulent mixing of heat, momentum, mass and constituents (Ralston and Stacey, 2005; Armenio and Sarker, 2002), and therefore limit the replenishment of dissolved oxygen in the bottom layer. Over time, benthic demand depletes dissolved oxygen to hypoxic levels. As a result, unlike many other coastal and estuarine systems, hypoxia in the Bay is thought to be naturally occurring and driven by stratification, instead of being man-made and driven by nutrient loadings.

Figure 5.1 shows the regions where stratification-induced hypoxia occurs in Corpus Christi Bay. The regions are delineated by. Paul Montagna, a marine biologist from the Harte Research Institute of Texas A&M University, Corpus Christi; and Ben Hodges, an environmental engineer at University of Texas at Austin.



Ernest Sin Chit To, CRWR

Figure 5.1. Hypoxia in Corpus Christi Bay

DESCRIPTION OF CORPUS CHRISTI BAY

The Corpus Christi Bay system is an urban estuary with complex hydrodynamic and water quality conditions (see Figure 5.1). The bay is approximately circular in shape with a diameter of approximately 13 miles. The average depth of the Bay is 9 feet, and is separated from the Gulf of Mexico by a barrier island, so that water circulation in the Bay is driven more by wind than by tides. Communication with the gulf is available via Aransas Pass (near Port Aransas) and Packery Channel.

Adjacent to Corpus Christi Bay are four bays: Redfish Bay, Laguna Madre, Nueces Bay, Oso Bay. The first two, Redfish Bay and Laguna Madre, are lagoonal systems that are characterized by:

1. Shallow depth. They are about 3 to 4 feet deep – therefore shallower than Corpus Christi Bay.
2. Limited freshwater inflow. Most inflows are from surrounding surface runoff during storm events
3. Dense aquatic vegetation (Montagna, 1993).

Due to the shallowness of Redfish Bay, Laguna Madre and Corpus Christi Bay, as well as their limited interaction with the gulf, elevated salinity levels are found during periods in the summer when evaporation is high and precipitation is low. The typical salinity in the Gulf of Mexico ranges from 10 to 30 psu (practical salinity units). During the summer, salinity in Corpus Christi Bay can reach as high as 50 psu, while salinity in upper Laguna Madre can reach as high as 70 psu.

The other two bays, Nueces Bay and Oso Bay are estuarine systems which are fed by freshwaters from the Nueces River and Oso Creek respectively. Because of this, Nueces Bay experiences lower salinities than Corpus Christi Bay. However in Oso Bay, hypersaline conditions can occur (up to 70 psu) because the Barney Davis Power Plant (see Figure 5.1) draws 400 million gallons a day of cooling water from the upper Laguna

Madre through an intake shown in the lower part of Figure 5.1, and discharges this flow into Oso Bay.

Gravity currents and hypoxia

In the previous chapter, hypersaline currents from Laguna Madre are shown to be the dominant cause of density stratification in southeast Corpus Christi Bay. Hypersaline currents are often referred to as gravity currents because their higher density (compared to the ambient fluid) causes gravity to control their movement. Gravity currents follow the bathymetry and travel towards deeper portions of the Bay. The emergence of gravity currents is controlled by strong southeasterly winds. The winds provide the hypersaline waters the momentum needed to overcome the resistance from the dense aquatic vegetation in Laguna Madre. The fate and transport of gravity currents are illustrated in Figure 5.2.

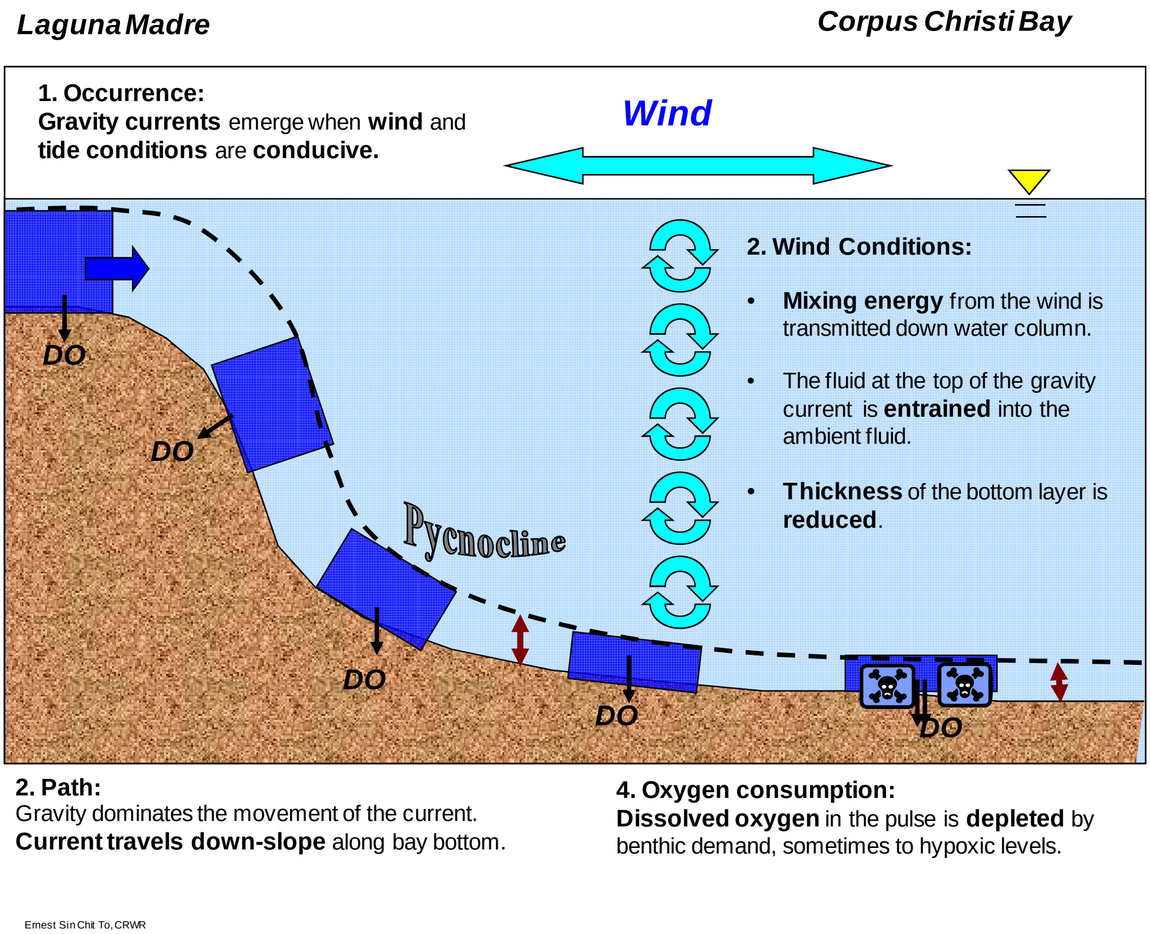


Figure 5.2. Gravity currents and hypoxia (To, 2008).

In summary, four conditions need to be satisfied in order for hypoxia to happen at a given location in the Bay. These are *occurrence, travel, wind conditions,* and *oxygen consumption.* First of all, a gravity current needs to emerge from the shallow bays. Secondly, the given location must be within the path of the current. Thirdly, wind conditions are not strong enough to break up the gravity current before it reaches the location. Lastly, dissolved oxygen is depleted below 2 mg/L when the gravity current reaches the location. An illustration is provided in Figure 5.3.

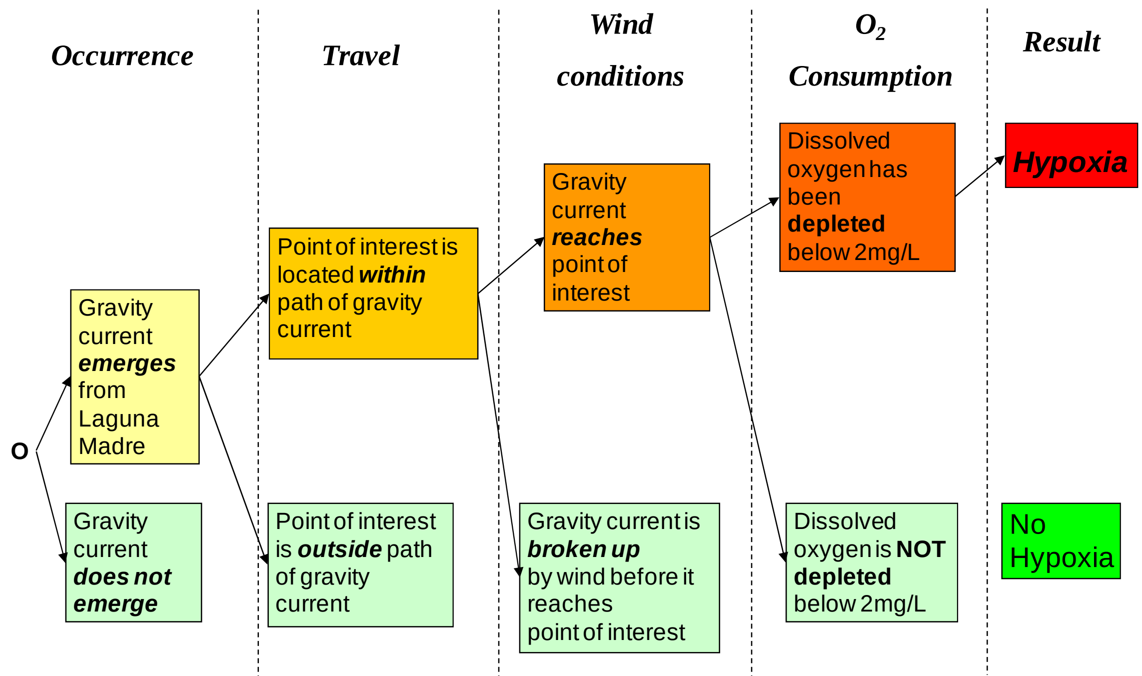


Figure 5.3. Conditions leading to hypoxia

DATA SOURCES IN CORPUS CHRISTI BAY

Several governmental agencies and academic institutions maintain environmental sensor networks in Corpus Christi Bay. These include the HRI (Harte Research Institute) network, Hodges surveys (funded by the Texas Water Development Board), TCOON (Texas Coastal Ocean Observation Network), and the TPWD (Texas Parks and Wildlife Department) biological sampling grid. The locations of the sensors are shown in Figure 5.4. A brief description of the data collected by these four networks is provided below.

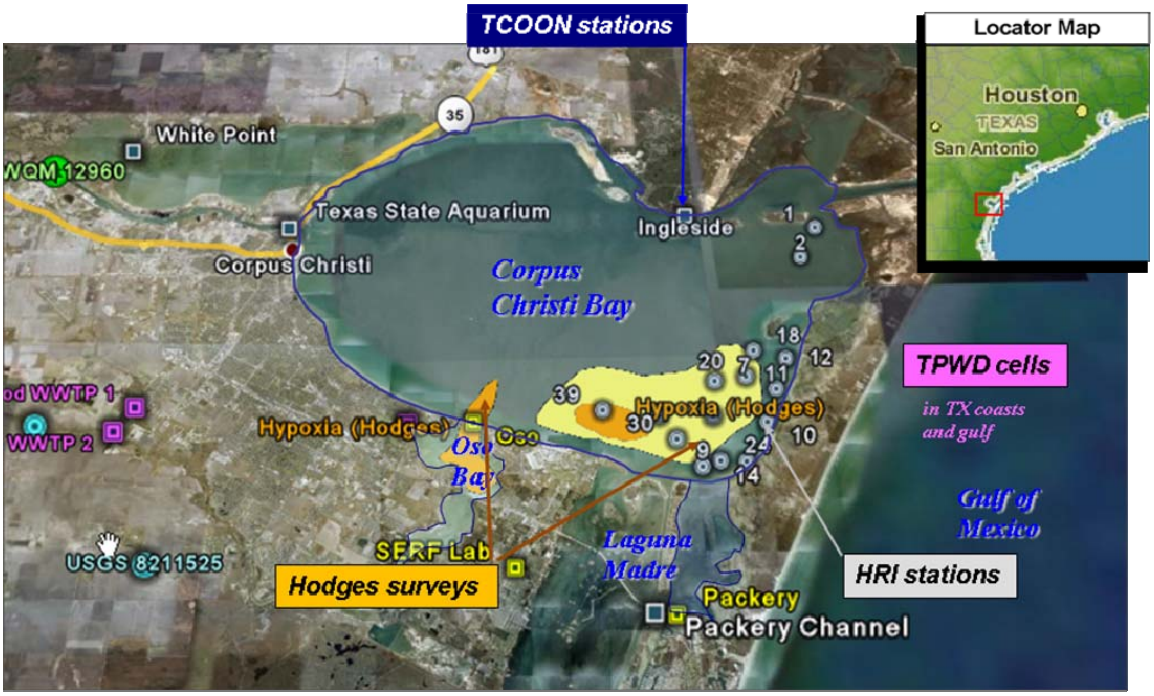


Figure 5.4. Sensor networks in Corpus Christi Bay.

Harte Research Institute (HRI) network

In east Corpus Christi Bay, Dr. Paul Montagna and his team have collected water quality data over a wide area (~ 50 square kilometers) since the 1980s. The monitoring stations of the network are shown in grey in Figure 5.4. Data are collected at weekly to bi-weekly intervals mostly during the summer months of June, July and August.

Hodges plume tracking studies

In the two regions downstream of Oso Bay and Laguna Madre, Dr. Ben Hodges and his team have performed <u>short-term</u> intensive data collection over smaller areas (see regions highlighted in orange in Figure 5.1) to capture the formation and dissipation of stratification in the Bay (Hodges and Furnans, 2006, Brower, *et al.* 2007, Hodges, et al., 2008). Salinity, oxygen and temperature data are collected every 12 hours. These high resolution studies have captured the emergence, movement and dissipation of gravity currents from Oso Bay and Laguna Madre.

Texas Coastal Ocean Observing Network

The Texas Coastal Ocean Observation Network (TCOON) is a set of observing stations located along the Texas coast that measure water level, wind speed, direction and a variety of environmental conditions, including water and air temperature. They collect data continuously at six-minute intervals. Six TCOON stations are located around Corpus Christi Bay, as shown in Figure 5.4. Historical data going back to 1993 are available for most parameters at these stations.

Texas Parks and Wildlife Department

The Texas Parks and Wildlife Department collects dissolved oxygen, salinity and temperature data in Corpus Christi Bay as part of its biological sampling program (see Figure 5.5). Sampling is performed based on a 1 minute by 1 minute grid along the entire Texas coast and its associated estuaries (TPWD, 2002). It is a spatially and temporally extensive data set that contains over 20,000 data points for CCBay, Upper Laguna Madre and Aransas Bay for the period from 1977 to present. Despite this, the temporal resolution TPWD data is limited. On average 2 samples are collected in each cell per year. Although these limitations prevent the data from capturing the fate and movement of gravity currents, the data are still useful for estimating the ambient conditions in Corpus Christi Bay and its adjacent bays.

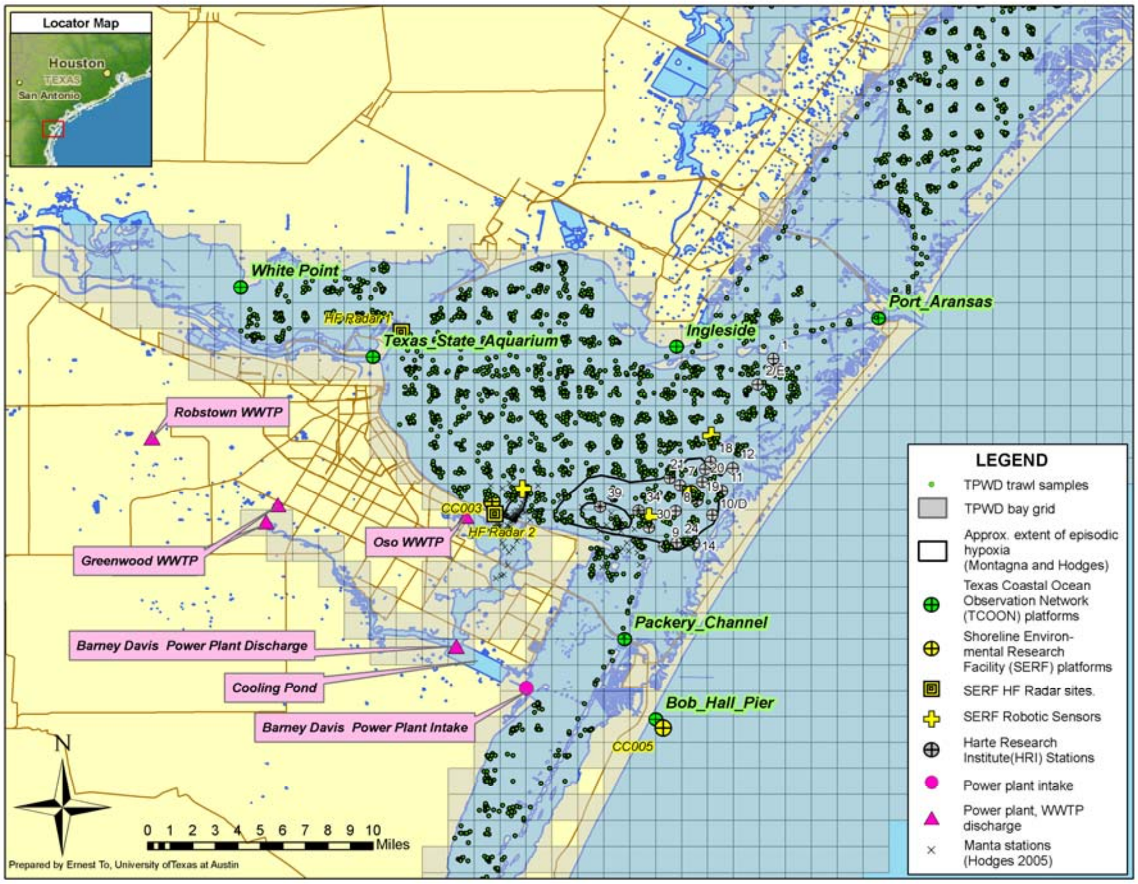


Figure 5.5. Location of water quality samples collected by the Texas Parks and Wildlife Department.

5.3 Using wind to model hypoxia patterns

CONCEPT OF THE VALVE HYPOXIA MODEL

The previous chapter demonstrated that winds of certain speed and direction provide the external force to push gravity currents into the Bay. The resulting mechanism is like operating a valve on a hose. Some winds turn the valve on and release a gravity current, other winds leave the valve off and release no gravity currents. A wind event is defined as a wind that turns the valve on. More wind events leads to stratification in a larger portion of the Bay. The valve hypoxia model extends this concept to explain the spatial and temporal pattern of hypoxia by tracking the fate and transport of gravity currents once they have been released into the Bay. The model is a simple plug flow model that takes into account oxygen depletion and the dissipation of the gravity current by the wind. To minimize the parameters needed to run the model, a parsimonious approach was taken. This meant making simplifying assumptions about gravity current physics and dissolved oxygen mechanics so that the model is workable with the sensor data available.

The model concept is described as follows (see Figure 5.6):

1. Wind events control the emergence of gravity currents similar to the operation of a valve.
2. The gravity currents travel along the bottom of the Bay towards lower elevation.
3. Some currents are eventually dissipated by wind mixing.
4. Some currents become hypoxic due to depletion by sediment oxygen demand.

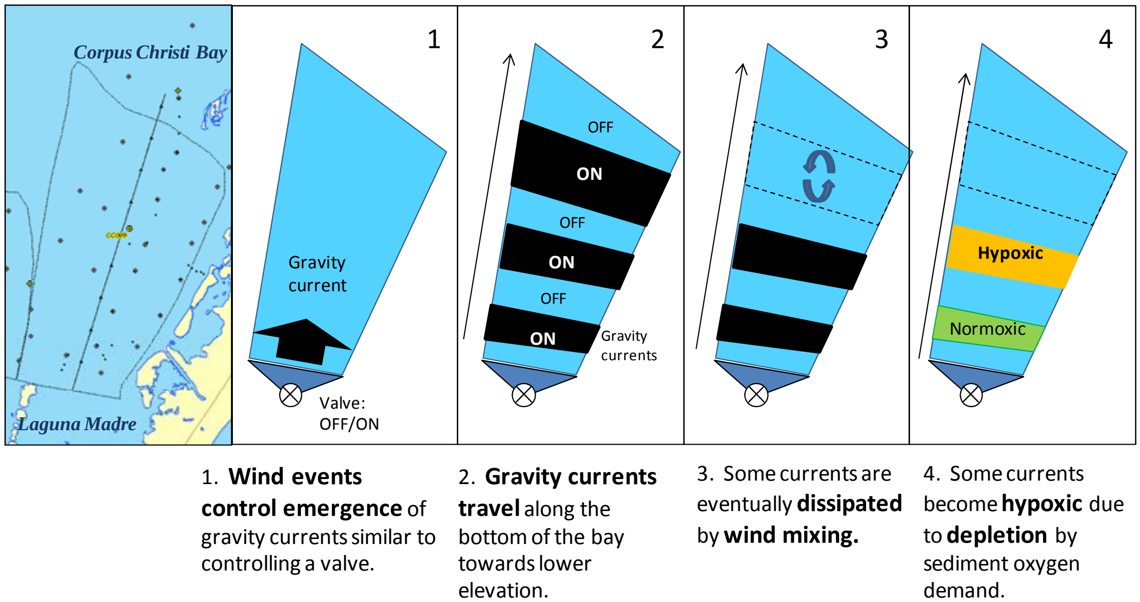


Figure 5.6. Concept of the valve hypoxia model

The model is Lagrangian because it tracks the fate and transport of every gravity current that emerges from Laguna Madre. The purpose of this model is to offer a basic framework for building future models to predict gravity-current-related hypoxia.

DOMAIN OF THE VALVE HYPOXIA MODEL

The model domain is is situated in southeast Corpus Christi Bay (see area outlined by dash line in Figure 5.7. This is an area where bathymetric features form an underwater basin that receives hypersaline waters from the eastern portion of Laguna Madre. As a result it presents an isolated system where one may observe gravity currents emerging from a single source (East Laguna Madre). This area is defined as the domain of the valve hypoxia model.

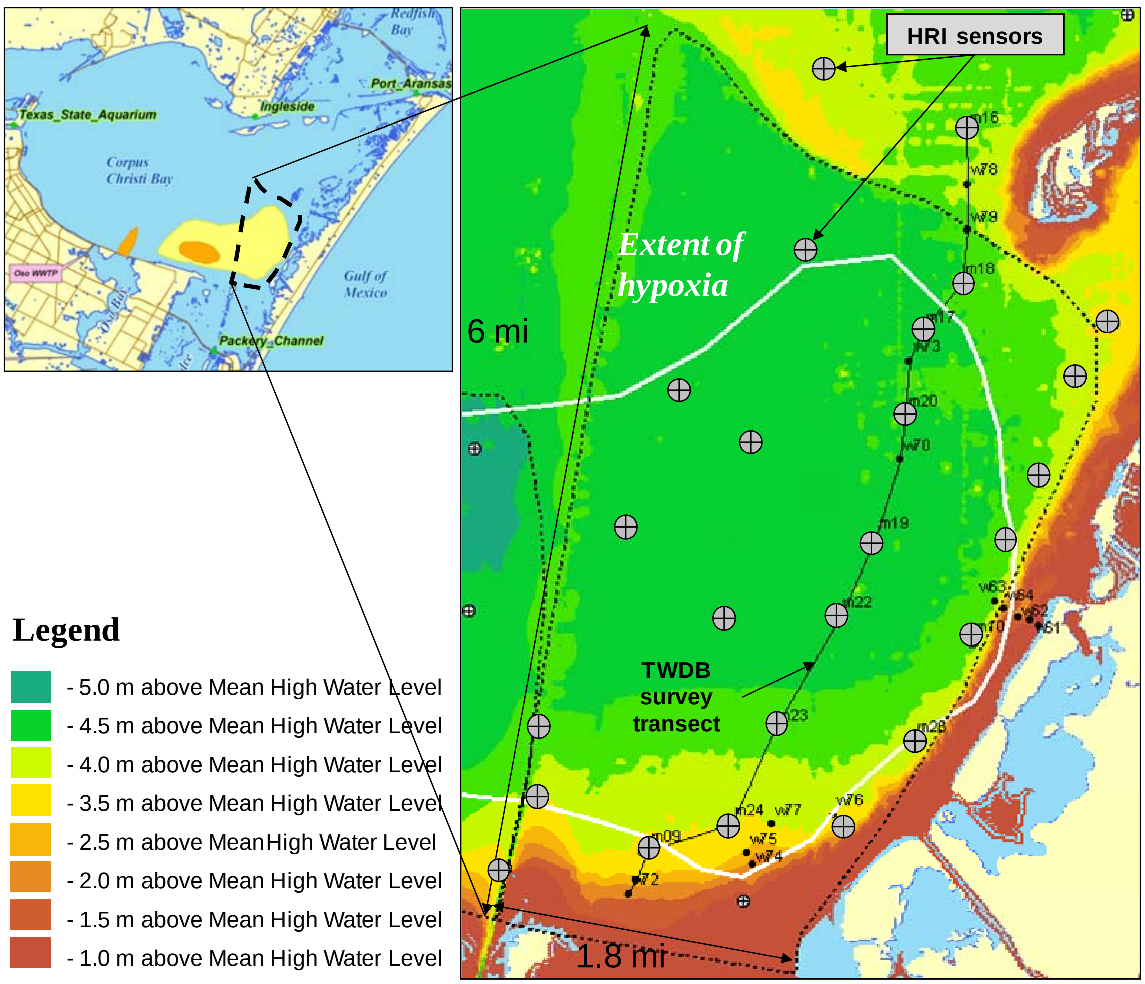


Figure 5.7. Domain of the valve hypoxia model.

ALGORITHM OF THE VALVE HYPOXIA MODEL

Create a time series of the valve state

Because every wind blowing across Corpus Christi Bay elicits an instantaneous response from the valve: ON or OFF, the time series of the wind vectors is converted into a time series of the valve state. Based on the wind-stratification analysis of the previous chapter, wind that blow from N120 ^{o}W to N165 ^{o}W at a velocity between 4 and 8 m/s turn the valve ON while other winds leave the valve OFF. As a demonstration, the wind history for the ten days prior to July 13, 2006 12:00 pm (from now on abbreviated as 7/13/2006 PM) shown in Figure 5.8 is converted to a ten-day time series of the valve state (see Figure 5.9). The period of ten days is chosen as a conservative measure. Even though a five-day travel time is observed from TWDB surveys, it is a value that was estimated visually and has inherent uncertainties. Therefore a period of ten days is chosen so that a longer wind history is considered.

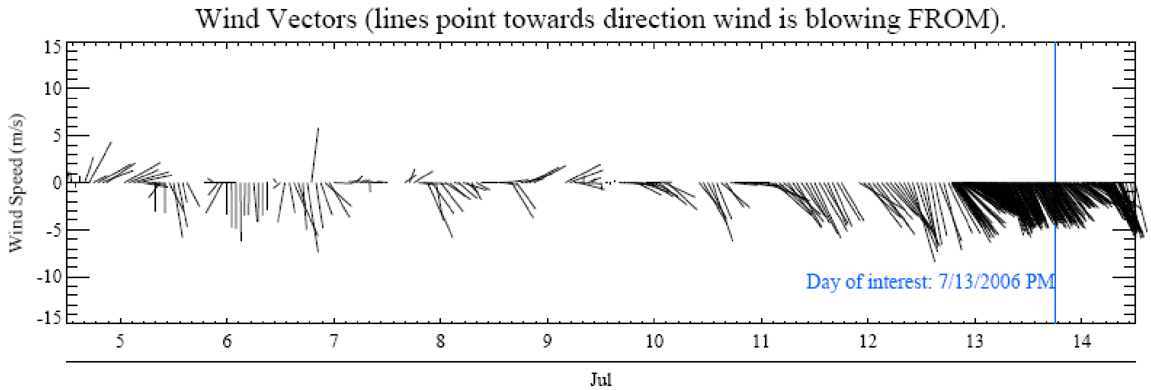


Figure 5.8	Time series of wind vectors in the ten days prior to July 13, 2006 afternoon.

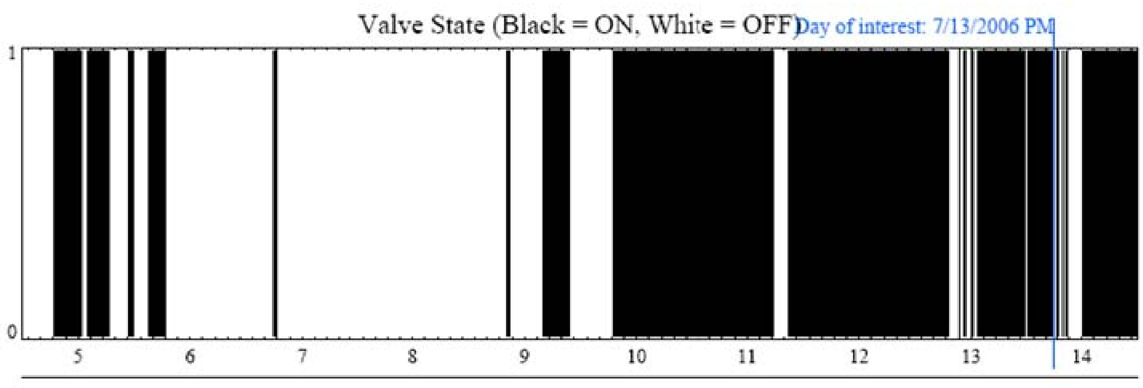


Figure 5.9	Time series of valve states in the ten days prior to July 13, 2006 afternoon.

The time series of the valve state indicates when in recent history have gravity currents been issued into the Bay. The isolation time is the length of time a gravity current has spent in the Bay. For example, a gravity current that was issued on 7/11/2006 6:00 a.m. would have an isolation time of 2.25 days on 7/13/2006 12:00 p.m.

Convert valve states to gravity current locations

The locations of gravity currents (i.e. described as distance from the mouth of Laguna Madre) are calculated using an assumed average speed multiplied by the isolation time. From the TWDB surveys, it is observed that gravity currents travelled at an average speed of ~1 km/day. Applying this speed to the gravity current issued on 7/11/2006 6:00 a.m. results in a computed location of 1 km/d x 2.25 d = 2.25 km from Laguna Madre on 7/13/2006 PM. If the entire ten-day wind history prior to 7/13/2006 PM is computed, then the location of all the gravity currents in the model domain (prior to accounting for wind mixing) can be obtained for the day of interest.

The spatial pattern of salinity along the depth and distance axes (prior to accounting for wind mixing) for 7/13/2006 PM is calculated by applying the following assumptions

1. Initial thickness of the gravity current is 1 m (based on observing data from the TWDB plume tracking study) ;
2. Salinity in Laguna Madre and Corpus Christi Bay are 40 ppt and 38 ppt, respectively (based on Texas Parks and Wildlife Data);

Figure 5.10 shows the preliminary salinity profile computed by the model.

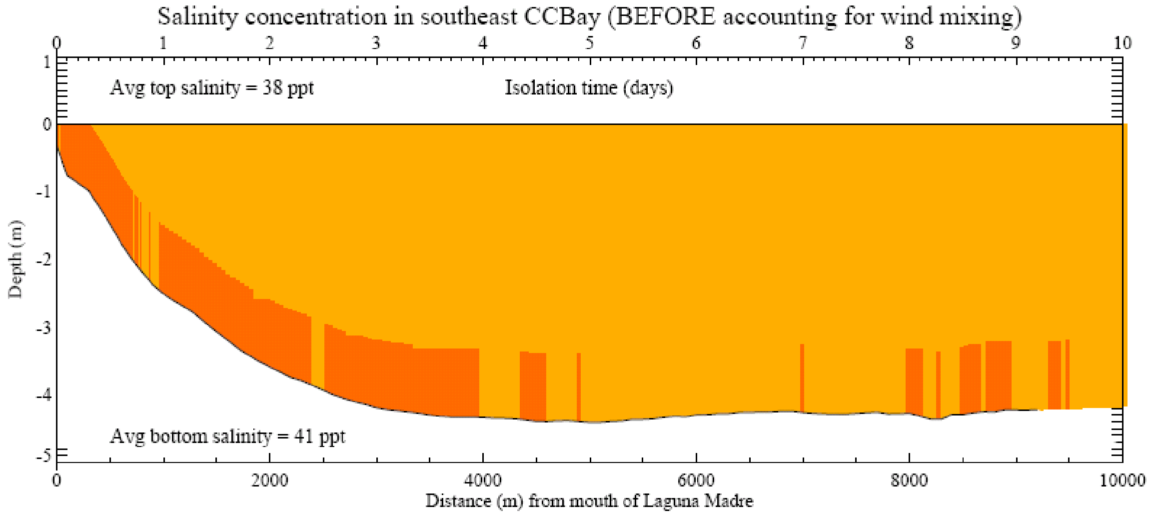


Figure 5.10. Salinity concentrations vs. depth and distance from mouth of Laguna Madre (before accounting for wind mixing).

Account for wind mixing

Equations from Hodges, *et al.* (2008) are used to quantify the effect of wind mixing on the thickness of the gravity currents. The energy transfer rate from the wind to per unit area of the water surface is quantified using Equation 5.1 (Hodges, et al., 2008):

$$\frac{de_w}{dt} \sim \frac{1}{2}\rho_a (C_N u_*)^3 \qquad \text{Equation 5.1}$$

where

C_N is an empirical coefficient ~ 1.4×10^{-3} (Hodges, *et al*., 2008);

u_* is the wind shear velocity; and,

ρ_a is the water density at the surface ~ 1.3 kg m^{-3}.

The wind shear velocity, u_*, is scaled on the wind speed, U_w, using Equation 5.2 (Hodges, et al., 2008):

$$u_* = U_w \sqrt{C_D \frac{\rho_{air}}{\rho_a}} \qquad \text{Equation 5.2}$$

where

ρ_{air} is the air density; and,

C_D is a drag coefficient ~ 1.33 (Hodges, *et al*., 2008).

Only a fraction of the wind-stirring energy is effectively used for mixing the gravity current because the majority is lost due to viscous dissipation. The rate of change of the mixing energy, e_p, is given by Equation 5.3 (Hodges, *et al*., 2008):

$$\frac{de_p}{dt} = C_m \frac{de_w}{dt} \qquad \text{Equation 5.3}$$

where C_m is the fraction of available kinetic energy that mixes the gravity current ~ 0.2 (Hodges, et al., 2008). Consider a gravity current that is issued at t_0 to t_1, integrating Equation 5.3 from t_0 to t_1 gives us the total energy accumulated, e_p, over the time period (see Equation 5.4) :

$$e_p = \int_{t_0}^{t_1} \frac{de_p}{dt} dt = \sum_{t_0}^{t_1} \frac{de_p}{dt} \Delta t \qquad \text{Equation 5.4}$$

where Δt is the time interval between wind observations.

By processing the entire ten-day wind history prior to July 13, 2006, the wind mixing energy accumulated in southeast Corpus Christi Bay can be obtained as a function of isolation time/distance (see Figure 5.11).

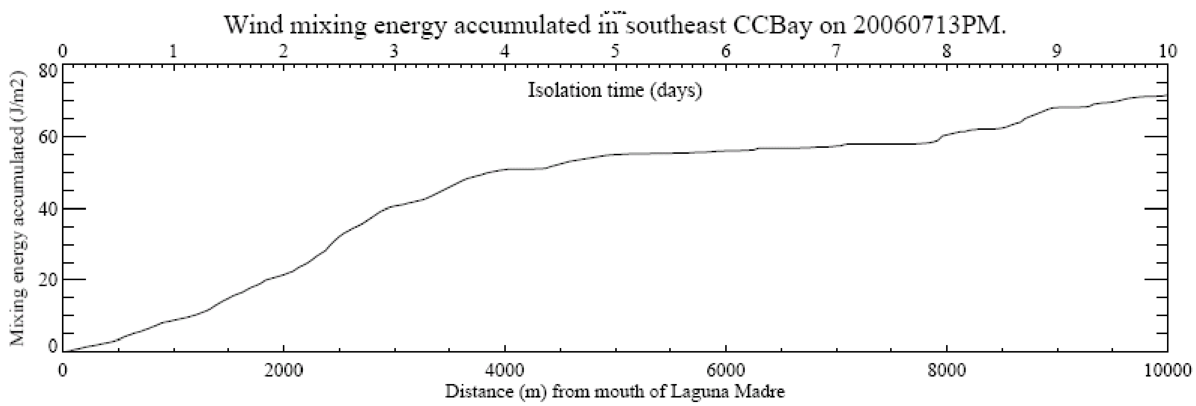


Figure 5.11. Wind mixing energy accumulated in Southeast Corpus Christi Bay on July 13, 2006 afternoon.

The mixing energy quantified in Equation 5.4 causes a top-down erosion of the dense underflow. Mixing increases the potential energy of the water column by pulling denser water from the bottom layer into the top layer. It is assumed that any dense fluid that is mixed is instantly evenly distributed throughout the water column of the top layer. The change in thickness of the bottom layer can be calculated by balancing the change in mixing energy and potential energy (see Equation 5.5 – Hodges, et al., 2008).

$$\frac{de_p}{dt} = -\frac{1}{2} g\Delta\rho(D - H_1)\frac{dH_1}{dt} \qquad \text{Equation 5.5}$$

where

$\Delta\rho$ is the density difference between the two layers.

D is the total depth of the water column.

H_1 is the thickness of the bottom layer that is eroded.

Two simplifying assumptions are made to ease the computation of the bottom thickness:

1) depth is constant with time (which is valid in the flat region of the Bay); and
2) $\Delta\rho$ does not change within period of interest because
    a. the top layer mixing laterally with adjacent waters and
    b. no ambient fluid is entrained into the bottom layer.

By incorporating these two assumptions and then integrating Equation 5.6 over the time interval, t_0 to t_1, Equation 5.6 is obtained:

$$0 = -\frac{H_{1t1}{}^2}{2} + DH_{1t1} - \left(\frac{e_p}{-\frac{1}{2}g\Delta\rho} - \frac{H_{1t0}{}^2}{2} + DH_{1t0}\right) \qquad \text{Equation 5.6}$$

where

H_{1t0} is the thickness of the bottom layer at t_0;

H_{1t1} is the thickness of the bottom layer at time t_1; and

e_p is the mixing energy that was accumulated between t_1 and t_0.

It can be seen that Equation 5.6 is quadratic with respect to the final thickness H_{1t1}. Solving Equation 5.6 results in two roots, where one is less than the water depth, D, and the other is greater than D. Because the thickness of the bottom layer can only decrease after mixing (thus $H_{1t1} \leq D$), only the lesser of the two roots is valid.

Equation 5.6 is used to compute the change in bottom thickness from after wind mixing to that before mixing. H_{1t0} and H_{1t1} referred to the thicknesses of the bottom layer before and after accounting for the wind.

As an example, let us consider a gravity current that is located at 2.25 km from the mouth of Laguna Madre and at a depth of 3.77m, with the following conditions:

1. thickness before considering mixing = H_{1t0} = 1 m
2. mixing energy accumulated over isolation time = e_p = 21.3 J/m^2
3. difference in top-bottom density of 2.3 g/L = 2.3 kg/m^3

Substituting these values into Equation 5.6, we get Equation 5.7:

$$0 = -\frac{{H_{1t1}}^2}{2} + 3.77H_{1t1} - \left(\frac{21.3}{-\frac{1}{2}(9.8)(2.3)} - \frac{1^2}{2} + 3.77 \cdot 1\right) \qquad \text{Equation 5.7}$$

Solving this equation yields two roots: 0.39 m and 7.15 m. Because the depth of the water is only 3.77 m, the only valid root is 0.39 m. Therefore the thickness of the gravity current after considering wind mixing is 0.39 m. By extending this computation process to the entire model domain, the salinity profile after accounting for wind can be obtained. Figure 5.12 shows the salinity profile in southeast Corpus Christi Bay on 7/13/2006 PM after considering wind mixing.

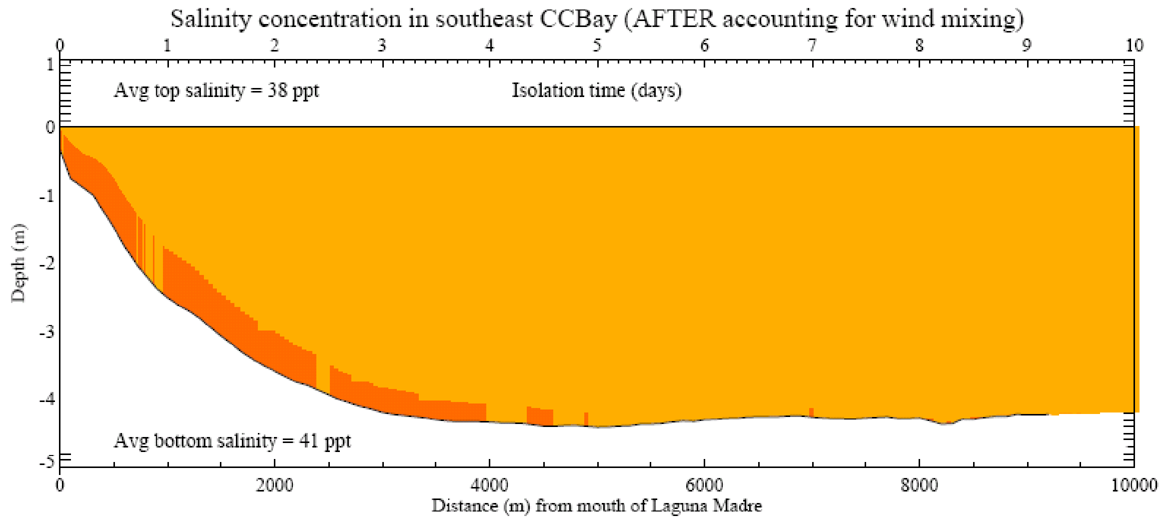


Figure 5.12. Salinity concentrations vs. depth and distance from mouth of Laguna Madre (after accounting for wind mixing).

Account for oxygen depletion

The oxygen depletion rate within the gravity current is assumed to deplete under a constant rate. A net oxygen demand rate of 0.18 mg/L/hr is used. This value is obtained from the average of the net oxygen demand rates published in Hodges, *et al.* 2008. It is also within the respiration rates stated in Russell and Montagna (2007), which ranges from 2 to 6 mg/L/d or 0.08 mg/L/hr to 0.24 mg/L/hr. Thus for a gravity current that was issued 2.25 days ago with an initial dissolved oxygen concentration of 9 mg/L, the current dissolved oxygen concentration would be 0 mg/L (9 mg/L – 2.25 days x 24 hrs/day x 0.18 mg/L/hr < 0 mg/L). Using this constant rate, the oxygen profile along the model domain is calculated and plotted. Figures 5.13 and 5.14 show the oxygen profile on 7/13/2006 PM before and after accounting for wind mixing and oxygen depletion respectively.

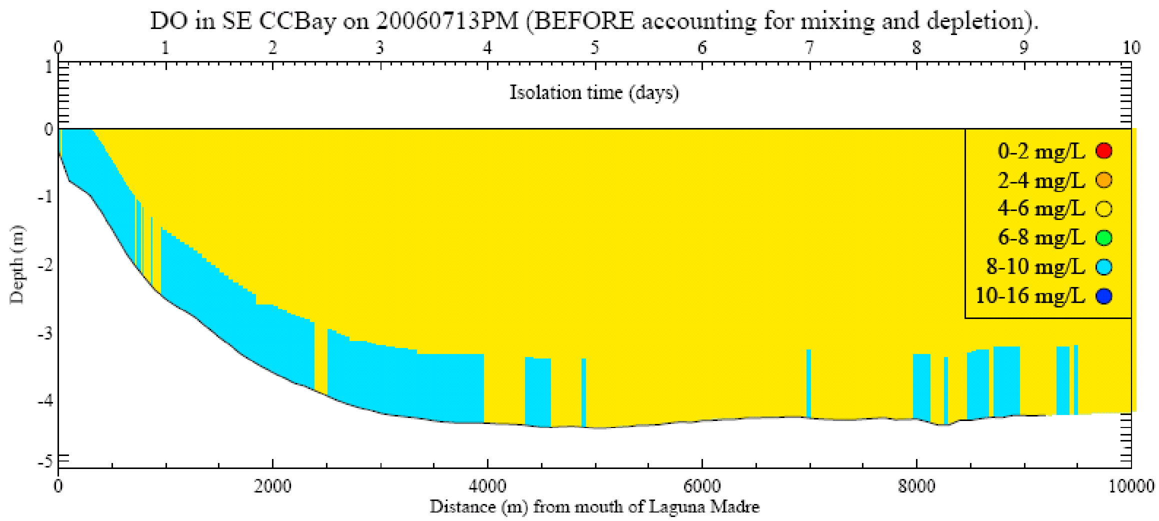


Figure 5.13 Salinity concentrations vs. depth and distance from mouth of Laguna Madre on 7/13/2006 PM (before accounting for wind mixing and oxygen depletion).

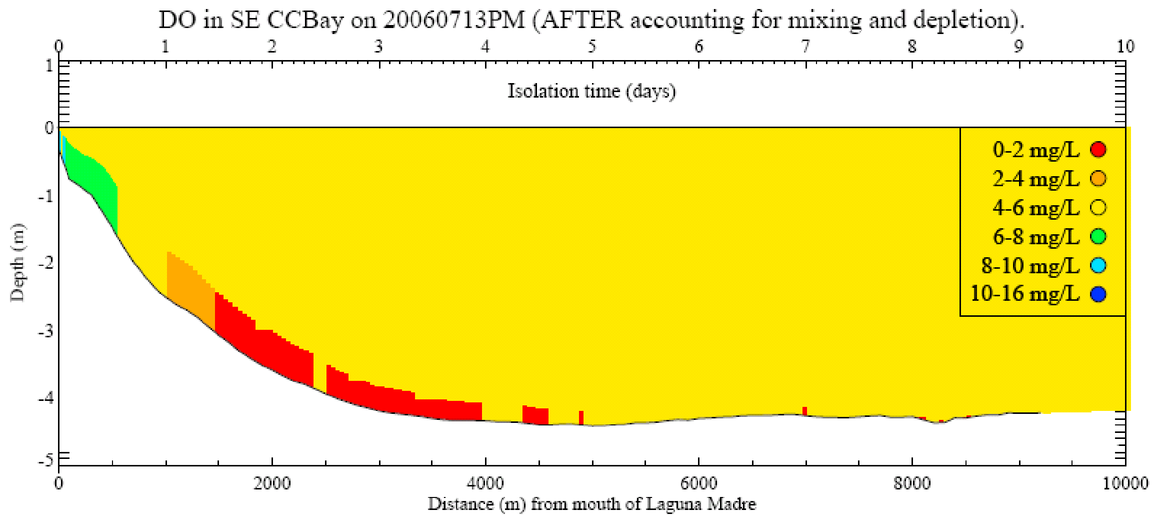


Figure 5.14 Salinity concentrations vs. depth and distance from mouth of Laguna Madre on 7/13/2006 PM (after accounting for wind mixing and oxygen depletion).

5.4 Results of the valve hypoxia model

MAPPING THE RESULTS

Results from the valve model are mapped for comparison with the salinity and oxygen profiles collected in the Bay. A simplifying assumption is made where gravity currents propagated forward in waves where fronts are perpendicular to the centerline of the experimental area. This is reasonable given the relatively rectangular geometry of the model domain (see Figure 5.15). The centerline of the model domain is depicted as a black line that runs from the mouth of Laguna Madre to the north. Distance markers are denoted as black triangles along the centerline. Salinity profiles collected on 7/13/2006

PM are represented as stacked charts on the map. The bottom salinity predicted by the model for 7/13/2006 PM (recall Figure 5.12) is shown as colored bands underneath the salinity profiles. Figure 5.16 shows the oxygen profiles and bottom oxygen concentrations predicted by the model for 7/13/2006 PM.

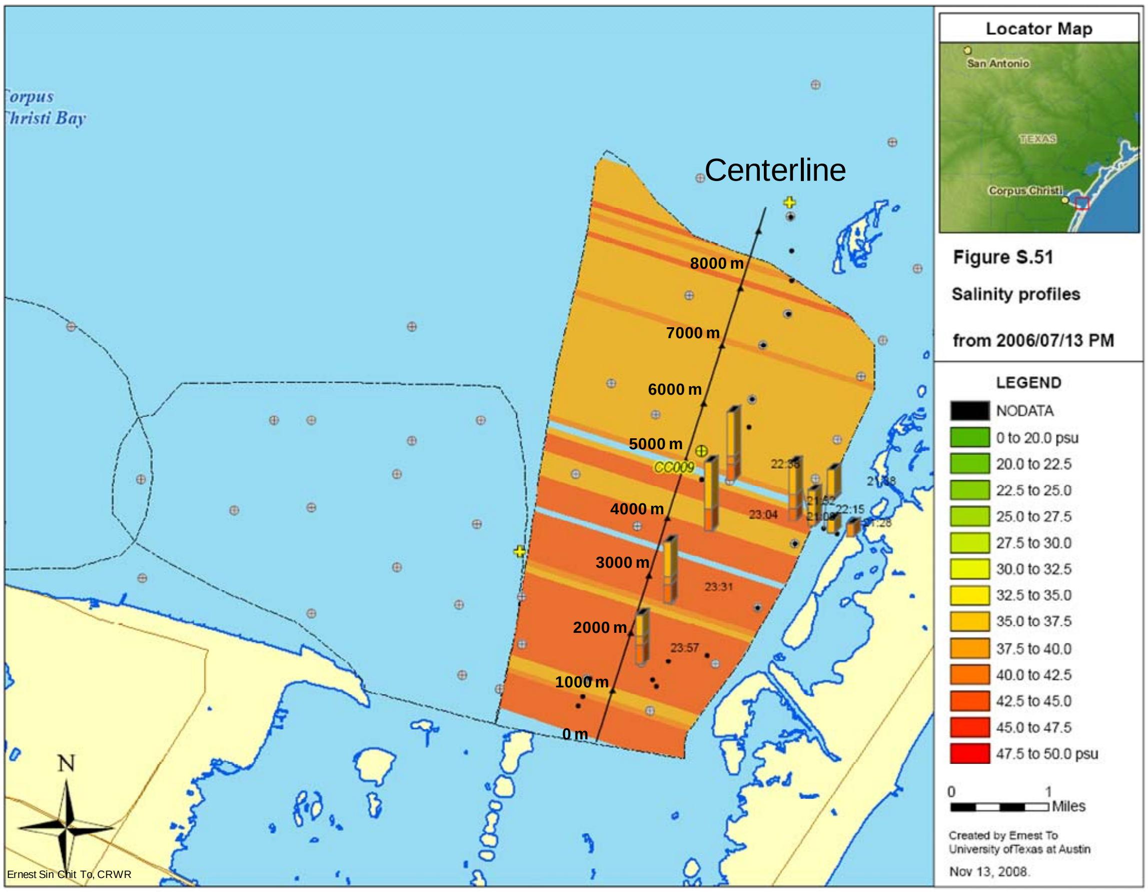


Figure 5.15 Bottom salinity predicted by the valve hypoxia model and salinity profiles collected on July 13, 2006.

The red zones in Figure 5.16 indicate where bottom oxygen concentrations are less than 2 mg/L. Therefore they indicate the hypoxic areas in the model domain.

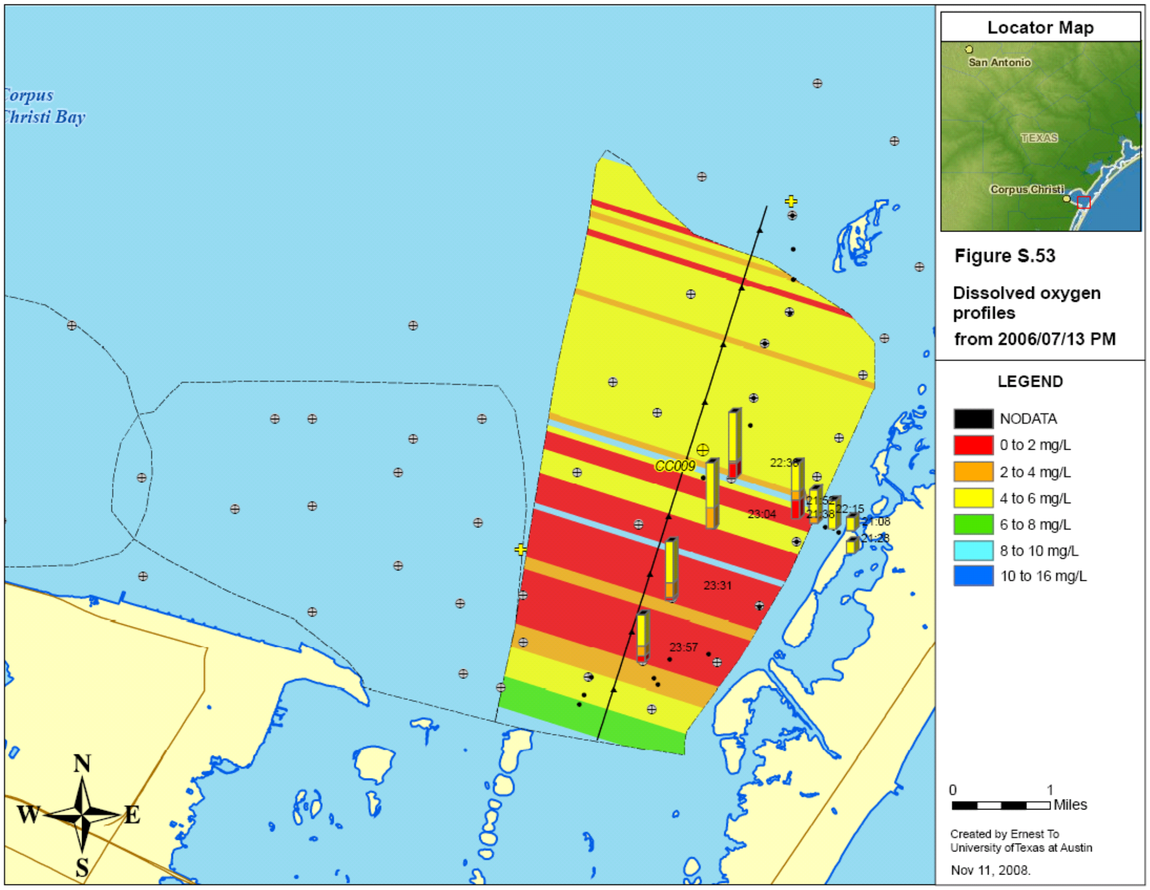


Figure 5.16 Bottom oxygen concentrations predicted by the valve hypoxia model and dissolved oxygen profiles collected on 7/13/2006 PM.

From Figure 5.15 and 5.16, it is observed that the general spatial distribution of gravity currents and hypoxic areas matched up well with the profiles collected in the Bay on 7/13/2006 PM.

RESULTS FROM 8/16/2005, 6/28/2005 AND 7/27/2005

The valve hypoxia model is also executed for the remainder of the 37 sampling days. Figures 5.17, 5.18 and 5.19 are maps of model results for the sampling days of 8/16/2005, 6/28/2005 and 7/27/2005. In each figure, predicted and observed salinities are presented on the left panel and predicted and observed oxygen concentrations are presented on the right panel.

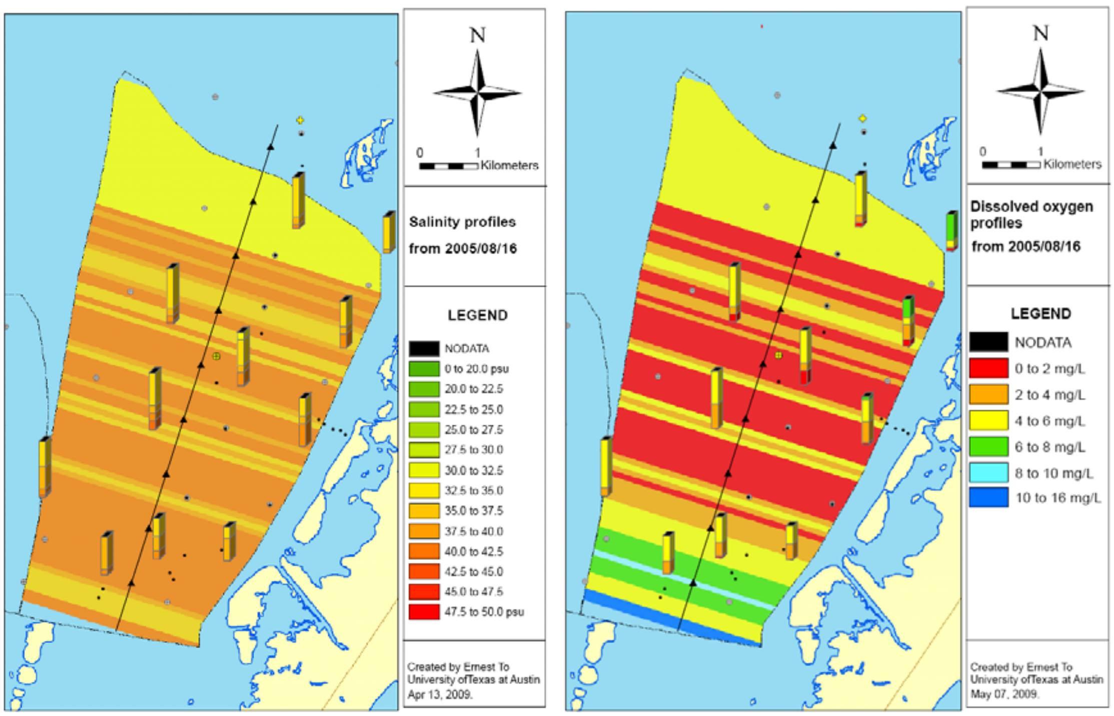


Figure 5.17. Maps of bottom salinity (left) and dissolved oxygen (right) predicted by the valve hypoxia model for 8/16/2005.

For 8/16/2005, the model predicts the presence of gravity currents in the majority of the model domain area. This agrees with the data because all the salinity profiles collected exhibits stratification. However, the model does not predict any stratification or hypoxia in the northern part of the Bay where the stratified profiles are found. Possible explanations include:

1. The gravity currents travel faster than the assumed speed of 1 km/d that is used in the model.
2. The gravity currents dissipate slower than predicted by the model.
3. The actual ON/OFF criteria for the valve model differ slightly from real conditions.

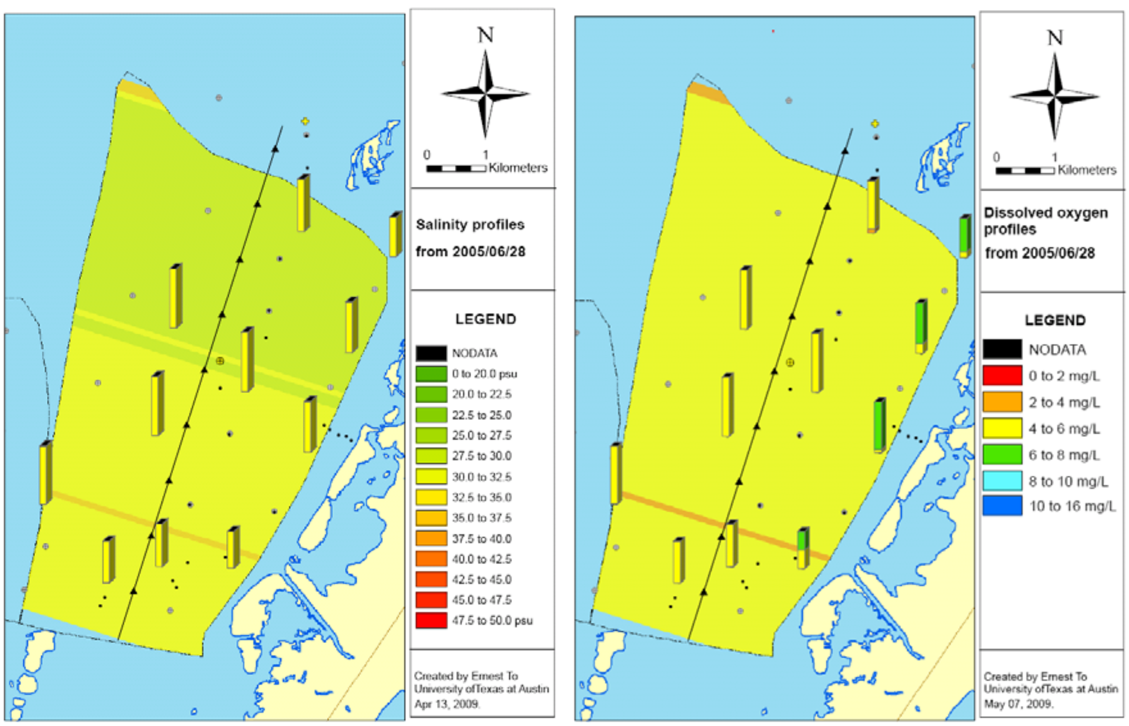


Figure 5.18. Maps of bottom salinity (left) and dissolved oxygen (right) predicted by the valve hypoxia model for 6/28/2005.

For 6/28/2005, the model predicts the absence of gravity currents from the majority of the model domain. This agrees with the data collected because none of the salinity profiles show stratification. For the oxygen profiles, it is observed that there is oxygen stratification on the east side of the model domain. However, the stratification is likely caused by highly-oxygenated freshwater running off Mustang Island in the east – which the model does not account for – and not by gravity currents from Laguna Madre. Nonetheless, the model's performance in predicting hypoxic areas is still reasonable.

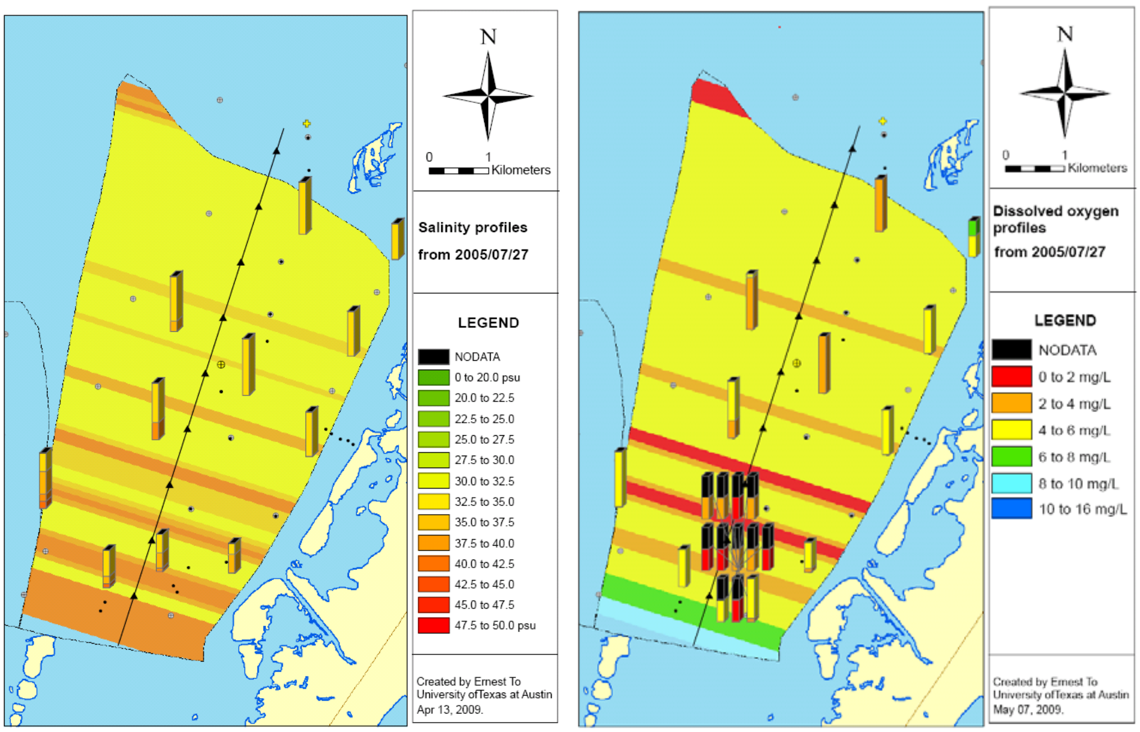


Figure 5.19. Maps of bottom salinity and dissolved oxygen predicted by the valve hypoxia model for 7/27/2005.

For 7/27/2005, the model predicts a concentration of gravity currents mostly in the south, with intermittent waves of currents in the middle region and almost no gravity currents in the north. This agrees with the general stratification pattern revealed by salinity profiles, although some discrepancies exist between the exact locations of stratification. This can be a result of the discrepancy in current speeds. For the oxygen patterns, the predicted and observed hypoxic areas are in general agreement with each other. As the model predicted, hypoxia is found between 2000 to 4000 m from the mouth of Laguna Madre. One item of interest is the cluster of oxygen profiles collected in the southern region of the domain (Figure 5.19 right panel). These profiles are data collected from continuous monitoring of dissolved oxygen at the Bay bottom over a 24-hour period. Each profile represents the average oxygen concentrations within a 2-hour interval. One can observe that actual oxygen concentrations fluctuated from hypoxic to marginally normoxic at different times of the day, even though they are all sampled at the same location. This phenomenon is the result of photosynthesis and respiration. Because the model does not account for diurnal fluctuations but uses an average depletion rate to compute changes in oxygen levels, it can under-predict oxygen concentrations during the day and over-predict during the night. Methods for improving the model in the future are described in section 5.6.

5.5 Implications of the valve hypoxia model

EXPLAINING HYPOXIA PATTERNS IN SOUTHEAST CORPUS CHRISTI BAY

Comparison between observations and model results (Figure 5.17, 5.18 and 5.19) showed that even for such a rudimentary model, the valve hypoxia model provides a logical explanation of salinity and oxygen patterns in the Bay. Data collected by HRI often find hypoxic areas interspersed with non-hypoxic areas in the Bay. This contradicts with the assumption that bottom salinity and hypoxia are continuous in space – as one would expect from *in situ* generation of stratification. The model demonstrates that wind-driven gravity currents, and therefore *ex situ* generation of stratification, are a much better explanation of hypoxia patterns in the Bay than *in situ* generation of stratification.

Conventional oceanographic wisdom suggests that shallow bays that have large surface areas like Corpus Christi Bay should be well-mixed during strong wind events. As such, stratification/hypoxia should only happen during extended calm wind periods. Contrary to this, the model shows that strong wind events can also cause stratification. In fact, the pattern of hypoxic vs. normoxic areas in the Bay can be related to waves of gravity currents flowing in as wind events turn the valve on and off.

The model results generally support the ON/OFF criteria deduced from wind stratification analysis. Several discrepancies between the predicted and actual gravity current patterns and hypoxia patterns are noted during the comparison. Methods for reducing these discrepancies are presented later in this chapter.

MANAGEMENT AND ENGINEERING IMPLICATIONS

With the understanding of *ex situ* stratification as the dominant cause of hypoxia patterns in Southeast Corpus Christi Bay, several management and engineering practices are suggested to control its occurrence:

1. On the upstream end, the source of hypersaline water can be mitigated by the following engineering measures.
    a. Dilute the salinity in Laguna Madre by building channels to the Gulf of Mexico. The main obstacle to this is that Laguna Madre is a large water body that extends all the way 150 km south from Corpus Christi Bay to Port Isabel. Several narrow inlets already exist along its length. The closest inlet is the Packery Channel. However, the tidal range provided by the inlets is relatively insignificant and has little effect on the hypersalinity in the lagoon. Significant dredging is needed to create a gap large enough to provide sufficient water for dilution.
    b. Build physical barriers to increase the wind speed threshold needed to push the gravity current into the Bay. One method is to construct low head dams at the mouth of Laguna Madre to prevent the down flow of gravity currents. The disadvantage is that such a dam prevents exchange between the Bay and Laguna Madre and exacerbates the hypersaline conditions in Laguna Madre. The presence of the dam can lead to conditions that are more catastrophic to the benthic life when the "dam" overtops during extreme wind events then without the dam.
2. On the downstream side, the gravity current can be mitigated by the following engineering measures:
    a. Enhance the dispersion the gravity currents in the Bay by dredging away terrain features that restrict bottom circulation; or,

b. Channel the gravity current to an even smaller area, for instance, into the ship channel, so that its effects are more localized and restricted.

5.6 Areas for improvement

The possible causes for discrepancies in the gravity current and hypoxia patterns are listed follows:

1. The ON/OFF criteria deduced for the valve model can differ slightly from actual conditions. This leads to discrepancies in the existence of gravity currents. The ON/OFF criteria can be refined in the future by repeating the wind-stratification analyses as more data are collected.
2. Discrepancies in gravity current speed.

   The constant speed assumption is an approximation of actual conditions. In reality, the speed of the gravity current varies depending on factors such as, the density difference between the gravity current and the ambient fluid, the slope of the Bay bottom and the lateral spreading of the current. Variations in speed lead to discrepancies in location of gravity currents. More sophisticated physical models can be incorporated in the future to better characterize gravity current movements.
3. Discrepancies in mixing mechanisms

   The mixing mechanisms govern how long gravity currents persist in the Bay. Two main assumptions are made to simplify the implementation of mixing dynamics. The first assumption is that the water column is approximated as a two box model where the density difference between the top and bottom layers is preserved throughout the mixing. Mixing only affects the thickness of the layers. In reality, entrainment of ambient fluid reduces the density difference between the two layers. The second assumption is that wind mixing occurs in a stepwise manner, where:

- In the first step, the model simulates the propagation of the gravity currents in the Bay without taking into account interactions with the ambient fluid. An unmixed salinity pattern in the Bay is computed based on valve states, travel time and an assumed initial thickness.

    - In the second step, the model introduces mixing energies that are accumulated over the isolation times of the gravity currents into the Bay. Physically, this is analogous to a sudden violent stirring of the water column that causes the salinity pattern to change instantaneously. The mixing energy changes the thickness of the gravity current from a pre-mixed state to the post-mix state.

In reality, both propagation and mixing processes happen simultaneously over the isolation time of a gravity current. However modeling these two processes simultaneously, calls for a more sophisticated model and is therefore recommended for future work.

4. Discrepancies in density stratification

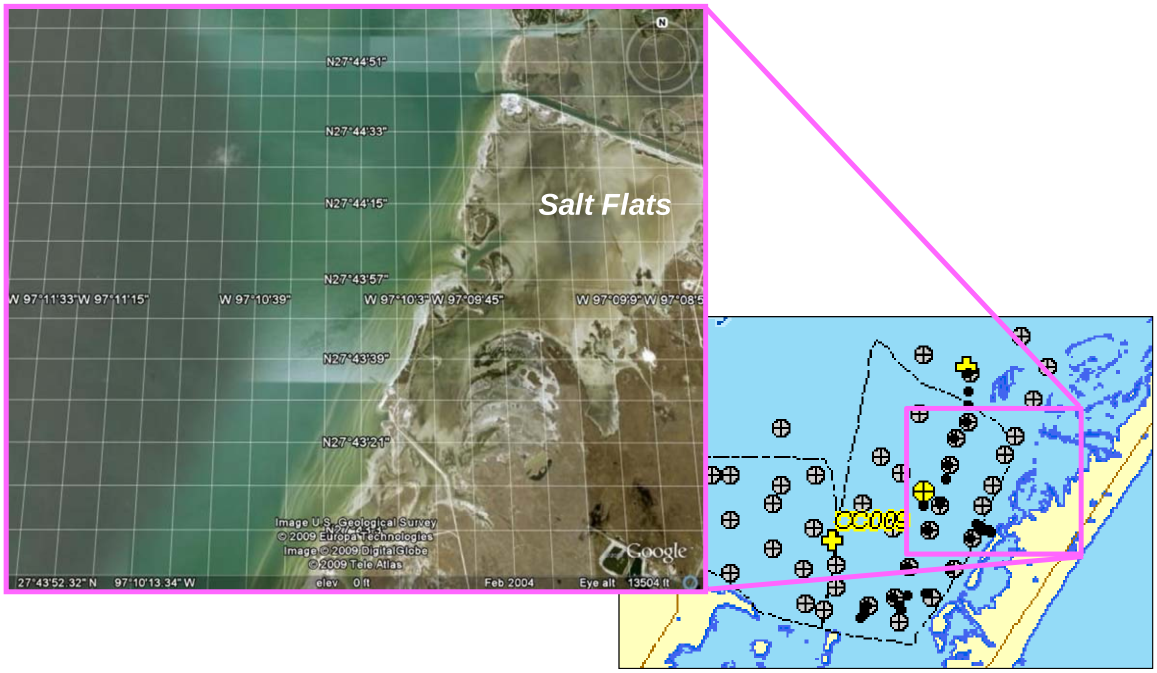
The mixing energy required to dissipate gravity currents is highly sensitive to the density difference between gravity current and ambient fluid. For instance, a density difference of 2 g/L between gravity current and ambient fluid requires twice as much mixing energy to dissipate than a gravity current with a density difference of 1 g/L. Therefore it is critical to obtain accurate estimates of salinities in Laguna Madre and Corpus Christi Bay. Unfortunately, TPWD data is the only source of water quality data available for this model. Because of the sparseness of its data, 90-day moving averages are used to interpolate estimates for the day of interest. This solution is less than ideal due to the fact that salinity in both Laguna Madre and Corpus Christi Bay can vary significantly after precipitation events. Higher sampling frequencies are needed to accurately characterize salinity. Recently, a real time sensor is installed by SERF (Shoreline Environmental Research Facility by Texas A&M University) in the center of the model domain to measure salinity in the Bay. However, no sensors are available to measure salinity in Laguna Madre.

5. Discrepancies in dissolved oxygen mechanisms

   The model assumes a constant rate of oxygen depletion while in reality, dissolved oxygen concentrations fluctuate diurnally because of photosynthesis and respiration. Better biological models can be used in the future to characterize dissolved oxygen dynamics more realistically.

6. Other sources of hypersaline flows

   The potential presence of other sources of hypersaline water in the northern part of the domain can cause discrepancies in the model results. For instance, salt flats are identified near Wilson's cut on Mustang Island (see Figure 5.20). These salt flats drain into the study area. However, because thes flats have not been sampled frequently, the extent of their influence is unknown. Because of their relatively small areas (compared to Laguna Madre) they are assumed by the model to have negligible influence. However, future studies need to be conducted in order to confirm this.



(Left image courtesy of Google Earth)

Figure 5.20. Presence of salt flats in the northeastern portion of the study area.

5.7 Conclusions for Chapter 5

This study demonstrated how wind-driven gravity currents can cause hypoxia in southeast Corpus Christi Bay. A simple plug flow model (also known as the valve hypoxia model) was constructed to simulate the transport and dissipation of gravity currents that emerge during these wind events. The model also simulated oxygen depletion within the currents. Results from the model demonstrated that the heterogeneous distribution of hypoxic areas in the Bay can be explained by of waves of hypersaline water flowing into the Bay during wind events. The model results also challenged the conventional oceanographic wisdom that shallow bays that have large surface areas like Corpus Christi Bay should be well-mixed during strong wind events. Contrary to this, the model shows that strong wind events can also cause stratification. The valve hypoxia model is presented as a framework for building models to predict wind-driven hypoxia. Its various simplifying assumptions can be replaced by more sophisticated models that better characterize gravity current movements and dissolved oxygen mechanics.

CHAPTER 6: CONCLUSIONS OF THE DISSERTATION

The framework proposed for using a Hydrologic Information System to study and predict hypoxia has been implemented successfully in Corpus Christi Bay. Methods for performing data compilation, data synthesis, hypothesis testing and predictive modeling of the hypoxia phenomenon are proposed and demonstrated. This chapter answers the four research questions of this dissertation, describes the contributions of this research, and provides recommendations for future work.

6.1 Answering the research questions

1. How can data be assembled from a service-oriented architecture of environmental sensor networks to describe the properties of a water domain in space and time?

Chapter 2 presents a methodology for harvesting and managing data from multiple sources. The methodology draws together a number of technologies (i.e. hydrologic information systems, web services, HydroGET, MySelect and ArcHydro) into one process. It is applied to the Corpus Christi Bay testbed project to help gather data to perform scientific research on hypoxia. The methodology is described as follows: First, a hydrological information system is constructed to provide unified access to environmental sensor networks via a service-oriented architecture. Next, the user creates a table called MySelect, which serves as a shopping list for data from sensor network and as a lookup table for mediating semantic differences among various networks. Third, the MySelect table is processed by HydroGET, a web service client that can harvest data from web services, to get the data from the HIS. Finally, HydroGET stores the data into an Arc Hydro data model, which utilizes the space-time data cube model to store time series data. While storing the data into Arc Hydro, HydroGET uses MySelect to group like variables together and separate unlike variables in the database. By utilizing this methodology, data can be consolidated from multiple sensor networks and organized into a format that is amenable to data analysis by researchers.

2. How can data that were collected at different locations, times, spatial resolutions and temporal frequencies be synthesized to provide a continual description of an environmental variable over space and time?

Chapter 3 demonstrates that space-time kriging can facilitate the construction of a continuous space-time domain of salinity from fragments of data collected by multiple sensor networks. De Cesare et al's modified GSLIB software proved to be a very useful tool for performing space-time kriging. The theory behind space-time kriging is largely similar to spatial kriging. However, the non-orthogonality among the spatial and temporal axes means that spatial and temporal variograms need to be modeled separately and then merged together using a product-sum or product model. Space-time kriging of salinity in Corpus Christi Bay from July 12 to 18, 2006 produces a continuous 3D space-time domain, where the dimensions are distance along a transect line, depth, and time. By dissecting this space-time volume in regular time intervals, snapshots of the gravity current as it undergoes stages of emergence, movement and dissipation are visualized. By interpreting the snapshots, the gravity current in the Bay is estimated to persist for a week and its average speed is estimated to be on the order of magnitude of 1 km/day.

3. How can data about different environmental variables be used to generate insight about underlying mechanisms about a given environmental phenomenon?

Chapter 4 presents a series of hypothesis tests for coupling wind and salinity conditions, whose results demonstrate that southeasterly winds blowing at velocities between 4 and 8 m/s are positively correlated with recorded instances of stratification. Moreover, the correlation between such winds and stratification is independent of the season they occur. Because wind is a much stronger factor than season in explaining stratification, the generation of stratification is dominated by the *ex situ* mechanism of gravity currents, rather than by the *in situ* mechanism of evaporation of bay waters.evaporation).

4. How can new models be designed around hydrologic information systems to make predictions about a given environmental phenomenon? What do models explain the hypoxia patterns in southeast Corpus Christi Bay?

In Chapter 5, insights from the wind-stratification analysis are incorporated into a plug flow model called the valve hypoxia model. The valve hypoxia model determines the release of gravity currents into the Bay based on the occurrence of strong, southeasterly winds in the recent wind history. The mechanism is analogous to operating a valve on a hose. Wind events turn the valve ON and release gravity currents. Non-wind events turn the valve OFF. The valve hypoxia model uses the gravity current speed, wind-mixing, and oxygen depletion mechanisms to predict the level of hypoxia. The valve hypoxia model was connected to real-time wind data from the TCOON web service and salinity and oxygen data from the TPWD web service to predict hypoxia patterns in the Bay.

Results from the model showed agreement with the general hypoxia patterns observed in the Bay. Therefore the model is able to explain the following about hypoxia phenomenon:

1. Hypoxia is caused by *ex situ* generation of stratification (i.e. by gravity currents from Laguna Madre) rather than *in situ* (i.e., by evaporation). Spatially heterogeneous hypoxic profiles can be explained as the result of waves of hypersaline water that emerge into the Bay during wind events.
2. Strong southeasterly winds that are faster than 4 m/s can force gravity currents into the Bay. Prior to this research, researchers have been perplexed about the occurrence of hypoxia after an extended period of stiff winds. Conventional oceanographic wisdom suggests that shallow bays that have large surface areas like Corpus Christi Bay should be well-mixed after strong wind events.

However field data, especially from the plume tracking studies contradict this wisdom. With the valve hypoxia model, the presence of stratification after strong wind events is explained as the result of strong southeasterly winds pushing gravity currents from Laguna Madre into Corpus Christi Bay.

6.2 Contributions to science and technology

The contributions of the proposed research to scientific knowledge are two-fold. To the field of collaborative environmental research, this research provides the following contributions:

1. Development of a framework for using HIS in environmental research which can be applied to a broad spectrum of environmental studies;
2. Creation of a methodology and set of tools (e.g. HydroGET, MySelect) for assembling data from a service-oriented architecture of environmental sensor networks to describe the properties of a water domain in space and time;
3. Demonstration of space-time kriging as a versatile method for synthesizing the data that are scattered over time and space, and is therefore an appropriate method for dealing with the data that are made available by HIS;
4. Demonstration of how data harvested from HIS can be used to build a parsimonious model to explain an environmental system; and, how the model, in turn, can be connected to real time data streams from HIS to predict the behavior of the environment; and,

5. An example of how HIS can be an integral component of collaborative environmental research.

To the investigation of the hypoxia phenomenon in Corpus Christi Bay this research contributes the following:

1. Evidence that extrinsic factors (e.g. gravity currents) is the dominant cause of stratification in bay and estuaries.
2. Visualization of the initiation, movement and dissipation of gravity currents in the Bay that allows estimation of the gravity current speed;
3. Proof that a positive relationship exists between southeasterly winds and stratification which enables the prediction of the onset and spatial pattern of stratification and hypoxia in southeast Corpus Christi Bay;
4. Creation of a simple valve hypoxia model that explains
    a. the heterogeneity in hypoxic patterns in southeast Corpus Christi Bay are caused by waves of gravity currents emerging from Laguna Madre;
    b. the occurrence of hypoxia typically observed after extended periods of strong winds from the southeast are caused by gravity currents pushed into the Bay.
5. Identification of data and knowledge gaps in the hypoxia phenomenon that can be covered by future investigation and sampling.

6.3 Next steps and recommendations

The following recommendations are proposed for the improvement of the framework for using HIS in scientific research:

1. Integrate data discovery with data compilation

   The methodology in Chapter 2 requires the user to create the MySelect table so as to specify what kind of data is harvested from each sensor network. This necessitates knowledge of sites and variables monitored by each network. To create a MySelect table, the user relies on other HIS applications like HydroExcel to get a catalog of each network. The catalog provides information on the availability of data for each site and variable and organizes the information in the MySelect format. Another HIS application that performs in data discovery is HydroSeek. HydroSeek is a map-based program provides users the ability to search for data across several networks using variable ontology by specifying general concepts such as nitrogen, oxygen, and streamflow. By integrating the capabilities of HydroGET, HydroSeek and HydroExcel together, a more streamlined process for data discovery and compilation can be achieved, which will greatly enhance the framework for scientific research.

2. Incorporate uncertainty in space-time interpolation

   Data collected by different sensors can contain different levels of uncertainty. One limitation of space-time kriging is that it treats data as discrete measurements without considering their errors. This results in less accurate synthesis results. Newer geostatistical methods such as the Bayesian Maximum Entropy (BME) method (Christakos, 2000) have the ability to incorporate both hard (discrete) and soft (probabilistic) data in its estimation and can be applied to improve the quality of data synthesis results.

3. Create an integrated analysis platform to perform all components of the scientific discovery framework.

   A variety of languages, ranging from Visual Basic, Fortran, MATBLAB, R to IDL; and applications, such as ArcGIS, Excel, and Google Earth were used in this research. To streamline the process of scientific discovery, an integrated analysis platform needs to be created to incorporate or access the functions of these programs. At the basic level, the platform can be a simple desktop application that has the basic functions of mapping, interfacing with HIS and some basic analytical functions. At the more advance level, the platform should have the ability to interoperate with other applications and to incorporate them into workflows. Existing applications such as ArcGIS already have strong capabilities to map, analyze, interact with web services and implement workflows. New applications such as NCSA's CyberIntegrator (Marini, *et al*., 2006) and the OpenMI Association's OpenMI (Moore, R.V., 2007) have the ability to incorporate programs created from disparate environments in analytical workflows. This platform may result from a mash-up of these different applications.

The following recommendations are proposed for the investigation of the hypoxia phenomenon in Corpus Christi Bay:

1. Increase real-time sensors in the Bay

   Real-time salinity and oxygen sensors in Laguna Madre and southeast Corpus Christi Bay are needed to accurately characterize the boundary conditions for the valve hypoxia model. The Texas Parks and Wildlife data are very sparse and the steps taken to interpolate the data mentioned in Chapter 5 probably reduced the model's accuracy. A real time sensor was installed by SERF (the Shoreline Environmental Research Facility of Texas A&M University) in southeast Corpus Christi Bay in August 2006, but had not been properly calibrated until recently. Another sensor is needed in Laguna Madre in order to properly characterize the upstream conditions for the model.

2. Perform more comprehensive hypoxia and plume tracking surveys

   The analysis for the valve hypoxia model was deduced from a data set spanning three years. Certainly with more hypoxia and plume tracking surveys, more accurate parameters can be derived (e.g. better ON/OFF criteria for the valve and gravity current speed). This can increase the accuracy of the predictive model.

3. Incorporate more sophisticated models to simulate gravity current speed and dissipation.

   The constant speed assumption used in the valve hypoxia model is an approximation of actual conditions. In reality, the speed of the gravity current can vary depending on factors such as density difference between gravity current and ambient fluid, ground slope and lateral spreading of the current. Variations in speed lead to discrepancies in location of gravity currents. More sophisticated physical models may be incorporated in the future to better characterize gravity current movements. The mixing mechanisms can be improved to take into account lateral entrainment of fluid into the gravity current. This will improve the accuracy in estimating the amount of wind energy needed to dissipate a gravity current. As a result, better predictions of the persistence of a gravity current in the Bay can be made.

4. Discrepancies in dissolved oxygen mechanisms

   The model assumes a constant rate of consumption while in reality, dissolved oxygen concentrations fluctuate diurnally because of photosynthesis and respiration. Better biological models can be used in the future to characterize dissolved oxygen dynamics more realistically.

CHAPTER 7: BIBLIOGRAPHY/ REFERENCES

Abdi, H, 2007, Bonferroni and Sidak corrections for multiple comparisons. in N.J. Salkind (ed.). Encyclopedia of Measurement and Statistics. Thousand Oaks, CA: Sage. http://www.utdallas.edu/~herve/Abdi-Bonferroni2007-pretty.pdf.

Applebaum, S., P.A. Montagna, and C. Ritter, 2005, Status and trends of dissolved oxygen in Corpus Christi Bay, Texas, U.S.A. Environmental Monitoring and Assessment 107: 297-311.

Armenio, V., S. Sarkar, 2002, An investigation of stably stratified turbulent channel flow using large-eddy simulation, Journal of fluid mechanics, 459, 1-42

Brower, A., Kulis, P., Hodges, B., 2007, Why can't they breathe? Wind Controlled Hypoxia and Stratification in Corpus Christi Bay. Presented at UT Austin ESI REU Poster Session.

Christakos, G., 2000, Modern Spatiotemporal Geostatistics, International Association for Mathematical Geology No. 6.

Christakos, G., Bogaert, P., Serre, M., 2002, Temporal GIS Advanced Functions for Field-Based Applications, Springer

Consortium of Universities for the Advancement of Hydrologic Science, Inc, 2007, Hydrologic Information System (HIS) Manual. Online at http://www.cuahsi.org/his/documentation.html

Coopersmith, E., Minsker, B., Maidment, D., Hodges, B., Bonner, J., Ojo, T, Montagna, P., 2007, An Environmental Information System for Hypoxia in Corpus Christi Bay: A WATERS Network Testbed, Congress Proceedings, World Environmental and Water Resources Congress 2007.

Cressie, N. and H-C Huang, 1999, Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions, Journal of American Statistical Association, Vol.94, No. 448. (Dec., 1999), pp. 1330-1340.

Dallimore, C., B.R. Hodges, J. Imberger, 2003, Coupling an underflow model to a 3D hydrodynamic model, Journal of Hydraulic Engineering, 129(10):748-757

Dallimore, C.J., J. Imberger and B.R. Hodges, 2004, Modeling a plunging underflow, Journal of Hydraulic Engineering, 130(11):1068-1076.

Dauer, D.M., A.J. Jr. Rodi, J. A. Ranasinghe 1992, Effects of low dissolved oxygen events on the macrobenthos of the lower Chesapeake Bay, Estuaries, 15, 384-391.

David, C. H., B.R. Hodges, 2006, Deploying a Microstructure Profiler in Corpus Christi Bay, CRWR Online Report, 2006, University of Texas at Austin.

De Cesare, L., D. Myers, D. Posa, 2001. Estimating and modeling space-time correlation structures. Statistics and Probability Letters 51 (1), 9–14.

De Cesare, L., D. Myers, D.Posa, 2002. FORTRAN programs in space-time modeling. Computers & Geosciences 28 (2002) 205–212.

De Iaco, S., D. Myers, D. Posa, 2001. Space-time analysis using a general product-sum model. Statistics and Probability Letters 52 (1), 21–28.

De Iaco, S., D. Myers, D. Posa, 2002. Nonseparable Space-time Covariance Models: Some Parametric Families. Mathematical Geology, Vol. 34, No. 1, January 2002.

Delhomme, J. P., 1978. Kriging in the HydroSciences. Advances in Water Resources Vol. 1, No. 5, p 251-266, September 1978.

Deutsch C. V., A. G. Journel, 1992. GSLIB: Geostatistical software library and user's guide. Oxford University Press, New York, 1992

Diaz, R.J. and R. Rosenberg, 1995, Marine benthic hypoxia: a review of its ecological effects and the behavioural responses of benthic macrofauna, Oceanography and Marine Biology Annual Review, 33, 245-303.

Diaz, R.J. and R. Rosenberg, 2008, Spreading dead zones and consequences for marine ecosystem, Science, 321, 926-931.

Englund, E. J., A Sparks, M. D. Robinson, 1999. Geo-EAS (Geostatistical Environmental Assessment Software). Environmental Software, 1989.

Goodall, J. L., J. S. Horsburgh, T. L. Whiteaker, D. R. Maidment, I. Zaslavsky, A first approach to web services for the National Water Information System, Environmental Modelling & Software 23 (2008) 404-411

Gray, J., R.S. Wu and Y.Y. Or, 2002, Effects of hypoxia and organic enrichment on the coastal marine environment, Mar. Ecol. Prog. Ser. 238 (2002), pp. 249–279.

Helsel, D. R., and Hirsch, R. M., 1995, Statistical methods in water resources, Elsevier Science Publishers, Chapter 15, pp 393-407.

Hodges, B., Furnans, J., and Kulis, P. 2008, A Thin-layer gravity current and hypoxia in Corpus Christi, submitted to Journal of Hydraulic Engineering; August 9, 2008.

Maidment, D. R.(ed.), 2002. Arc Hydro GIS for water resources, ESRI press.

Martin, C.M. and P.A. Montagna, 1995, Environmental assessment of LaQuinta Channel, Corpus Christi Bay, Texas, Texas Journal of Science, 47, 203-222.

Montagna P., 1993, Comparison of ecosystem structure and function of created and natural seagrass habitats in Laguna Madre, Texas., The University of Texas Marine Science Institute Technical Report Number TR/93-007.

Montagna P., 2007, verbal communication.

Montagna, P.A. and R.D. Kalke, 1992, The effect of freshwater inflow on meiofaunal and macrofaunal populations in the Guadalupe and Nueces Estuaries, Texas, Estuaries, 15, 307-326.

National Oceanic and Atmospheric Administration, 2007b, Short-Term Inundation Forecasting for Tsunami, http://www.ngdc.noaa.gov/

Moore, R.V., 2007, The OpenMI Association Strategy Statement for the next decade , 2nd OpenMI Life Workshop and Associated Meetings, 20-21 November 2007, Wallingford, UK

Marini L., B. Minsker, R. Kooper, J. Myers and P. Bajcsy, 2006, CyberIntegrator: A Highly Interactive Problem Solving Environment to Support Environmental Observatories, poster, EOS Trans. AGU, 87(52), Fall Meet. Suppl.., Abstract IN43B-0908, San Francisco, CA, December 11-15, 2006.

National Oceanic and Atmospheric Administration's (NOAA), 2007a, Hypoxia and Nutrient Pollution Overview, online at http://www.cop.noaa.gov/stressors/pollution/overview.html.

National Science and Technology Council, 2003, An Assessment of Coastal Hypoxia and Eutrophication in U.S. Waters. Online at http://oceanservice.noaa.gov/outreach/pdfs/coastalhypoxia.pdf.

Pearson, T. and R. Rosenberg, 1978, Macrobenthic succession in relation to organic enrichment and pollution of the marine environment, Oceanogr. Mar. Biol. Annu. Rev. 16 (1978), pp. 229–311.

Ralston, D.K. and M.T. Stacey, 2005, Stratification and turbulence in subtidal channels through intertidal mudflats, Journal of Geophysical Research, Vol. 110, No. C8, Ritter, C. and P.A. Montagna, 1999, Seasonal hypoxia and models of benthic response in a Texas bay, Estuaries, 22, 7-20.

Rouhani, S. and D. E. Myers, 1990, Problems in Space-Time Kriging of Geohydrological Data, Mathematical Geology, Vol. 22, No. 5, 1990.

Russell, M. and P. Montagna, 2007, Spatial and Temporal Variability and Drivers of Net Ecosystem Metabolism in Western Gulf of Mexico Estuaries, Estuaries and Coasts Vol. 30, No. 1, p. 137–153.

Russell, M.J., P.A. Montagna and R.D. Kalke, 2006, The Effect of Freshwater Inflow on Net Ecosystem Metabolism in Lavaca Bay, Texas, Estuarine, Coastal, and Shelf Science. Volume 68, Issues 1-2, June 2006, Pages 231-244

Simpson, John, E., 1999, Gravity Currents: In the Environment and the Laboratory Modeling, 2nd Edition, Cambridge University Press, ISBN 0521664012, 9780521664011, pp. 1.

Srinivasan, S, 2007, Stochastic Methods for Reservoir Modeling, lecture notes.

Texas Parks and Wildlife Department, 2002. Marine Resource Monitoring Operations Manual. Available on request from Texas Parks and Wildlife Department.

To, E. S. C, T. L. Whiteaker, E. S. Hersh, 2008, HydroGET – a web services client for Arc GIS. Available from http://his.cuahsi.org/hydroget.html (Accessed 12 Dec 2008).

To, E. S. C., Hypoxia modeling in Corpus Christi Bay using a Hydrologic Information System, PhD Dissertation Proposal, 2008, University of Texas at Austin.

To, E. S.C., and D. R. Maidment, 2008, Synthesis of web service data to support hypoxia modeling in Corpus Christi Bay. Presented at American Water Resources Association Spring Specialty conference, March 2008.

UNESCO, 1980, International State Equation for Seawater. http://unesdoc.unesco.org/images/0004/000473/047363eb.pdf

United States Environmental Protection Agency, 2008, Marine water quality indicators, http://www.epa.gov/region10/psgb/indicators/marine_wq/index.htm

Ward, G.H. and N.E. Armstrong, 1997, Current Status and Historical Trends of Ambient Water, Sediment, Fish and Shellfish Tissue Quality in the CCBNEP Study Area Technical Report CCBNEP-13, Corpus Christi Bay National Estuary Program, 280p. http://www.cbbep.org/publications/virtuallibrary/CC13A.pdf

Zar, J. H., 1999, Biostatistical Analysis, Prentice-Hall, Inc. ISBN 0-13-081542-X

VITA

Sin Chit To ("Ernest") was born in Hong Kong. In his youth he lived in Hong Kong and Singapore. He is a citizen of the latter country. In 1996, he obtained a Bachelor's degree in Civil Engineering from the University of Texas at Austin, after which he served in the Singapore Armed Forces as an infantry lieutenant. In 1999, he enrolled at Cornell University and graduated with a master's degree in Civil and Environmental Engineering in 2001. He then worked as an environmental consultant for Quantitative Environmental Analysis, LLC for more than four years before returning to his alma mater, University of Texas, to study for a doctorate.

Permanent address: 89B Tenth St, Hong Lok Yuen, Tai Po, N.T., Hong Kong, PRC.

This dissertation was typed by the author.